

## SEARCH REQUEST FORM

Requestor's  
Name: \_\_\_\_\_

Serial  
Number: \_\_\_\_\_

Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

### Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

### STAFF USE ONLY

Date completed: 08-27-03

Searcher: Beverly C 4994

Terminal time: 20

Elapsed time: \_\_\_\_\_

CPU time: \_\_\_\_\_

Total time: 23

Number of Searches: \_\_\_\_\_

Number of Databases: 1

#### Search Site

\_\_\_\_\_ STIC

\_\_\_\_\_ CM-1

\_\_\_\_\_ Pre-S

#### Type of Search

\_\_\_\_\_ N.A. Sequence

\_\_\_\_\_ A.A. Sequence

\_\_\_\_\_ Structure

\_\_\_\_\_ Bibliographic

#### Vendors

\_\_\_\_\_ IG

\_\_\_\_\_ STN

\_\_\_\_\_ Dialog

\_\_\_\_\_ APS

\_\_\_\_\_ Geninfo

\_\_\_\_\_ SDC

\_\_\_\_\_ DARC/Questel

✓ Other CGN

**THIS PAGE BLANK (USPTO)**



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 102216

To: Minh-Tam Davis  
Location: CM1-8E12  
Art Unit: 1642  
Wednesday, August 27, 2003

Cas Serial Number: 09/700700

From: Beverly Shears  
Location: Biotech-Chem Library  
CM1-1E05  
Phone: 308-4994

beverly.shears@uspto.gov

### Search Notes

priority date:  
05/21/98

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2003, 12:08:16 ; Search time 9197 Seconds  
(without alignments)  
8611.617 Million cell updates/sec

Title: US-09-700-700-1  
Perfect score: 1936  
Sequence: 1 aatggtatgccaaacttaagt.....ggggcgccgcgactagtga 1936

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg\_hum.\*
- 31: em.htg\_inv.\*
- 32: em.htg\_other.\*
- 33: em.htg\_mus.\*
- 34: em.htg\_pln.\*
- 35: em.htg\_rnd.\*
- 36: em.htg\_rnd.\*
- 37: em.htg\_vrt.\*
- 38: em.sy.\*
- 39: em.htgo\_hum.\*
- 40: em.htgo\_mus.\*
- 41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1836.4	94.9	2984	6	AR237420	AR237420 Sequence
2	1836.4	94.9	2984	6	AR260913	AR260913 Sequence
3	1836.4	94.9	2984	6	AR278444	AR278444 Sequence
4	1836.4	94.9	2984	6	AX106211	AX106211 Sequence
5	1836.4	94.9	2984	6	AX106554	AX106554 Sequence
6	1836.4	94.9	2984	6	AX140845	AX140845 Sequence
7	1836.4	94.9	2984	6	AX200705	AX200705 Sequence
8	1836.4	94.9	2984	6	AX267361	AX267361 Sequence
9	1834.2	94.7	80586	9	AC012574	AC012574 Homo sapi
10	1834.2	94.7	174445	9	AC051642	AC051642 Homo sapi
11	1832.6	94.7	184361	2	AC022597	AC022597 Homo sapi
12	1645.6	85.0	3266	9	AF247704	AF247704 Homo sapi
13	1388.8	71.7	2051	6	AR244306	AR244306 Sequence
14	641.2	33.1	1024	6	AR244053	AR244053 Sequence
15	628.4	32.5	1024	6	AR244068	AR244068 Sequence
16	628.2	32.4	1024	6	AR244054	AR244054 Sequence
17	622	32.1	1024	6	AR244067	AR244067 Sequence
18	615.6	31.8	1013	6	AR243986	AR243986 Sequence
19	611.8	31.6	724	6	AX203340	AX203340 Sequence
20	584.6	30.2	948	6	AR243995	AR243995 Sequence
21	582.2	30.1	720	6	AR244291	AR244291 Sequence
22	581	30.0	744	6	AR244290	AR244290 Sequence
23	580.6	30.0	716	6	AX203323	AX203323 Sequence
24	544.4	28.1	663	6	AX203318	AX203318 Sequence
25	539.6	27.9	980	6	AR243991	AR243991 Sequence
26	488	25.2	691	6	AX203331	AX203331 Sequence
27	477	24.6	697	6	AX203332	AX203332 Sequence
28	476.8	24.6	579	6	AX203314	AX203314 Sequence
29	469.4	24.2	484	6	AR261002	AR261002 Sequence
30	469.4	24.2	484	6	AR278533	AR278533 Sequence
31	469.4	24.2	484	6	AX106653	AX106653 Sequence
32	469.4	24.2	484	6	AX140944	AX140944 Sequence
33	469.4	24.2	484	6	AX200804	AX200804 Sequence
34	469.4	24.2	484	6	AX267460	AX267460 Sequence
35	461.6	23.8	582	6	AX203333	AX203333 Sequence
36	447.4	23.1	553	6	AX203313	AX203313 Sequence
37	446.6	23.1	683	6	AX203345	AX203345 Sequence
38	426.2	22.0	731	6	AX203310	AX203310 Sequence
39	421.6	21.8	1020	6	AR244074	AR244074 Sequence
40	421.6	21.8	1021	6	AR244075	AR244075 Sequence
41	411	21.2	822	6	AR244195	AR244195 Sequence
42	378.2	19.5	497	6	AR244194	AR244194 Sequence
43	358.8	18.5	374	6	AX337002	AX337002 Sequence
44	358.8	18.5	374	11	G37241	G37241 SHGC-57295
45	357.6	18.5	375	6	BD121246	BD121246 EST and e

ALIGNMENTS

RESULT 1	AR237420	2984 bp	DNA	linear	PAT 20-DEC-2002
LOCUS	Sequence 335 from patent US 6465611.				
DEFINITION	AR237420				
ACCESSION	AR237420				
VERSION	AR237420.1	GI:27282078			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 2984)				
AUTHORS	Xu,J., Dillon,D.C. and Mitcham,J.L.				
TITLE	Compounds for immunotherapy of prostate cancer and methods for their use				
JOURNAL	Patent: US 6465611-A 335 15-OCT-2002;				



SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2984)  
AUTHORS Mashiki, Z. and Harada, J.  
TITLE Negative pressure control apparatus for engine mounted in vehicle  
JOURNAL Patent: US 6321716-A 335 27-NOV-2001;  
FEATURES Location/Qualifiers  
source 1. 2984  
BASE COUNT 837 a 659 c 645 g 842 t 1 others  
ORIGIN

Query Match 94.9%; Score 1836.4; DB 6; Length 2984;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1897; Conservative 0; Mismatches 12; Indels 7; Gaps 5;

QY 1 AATGGTATGCCAACTTAAGTATTTACAGGGTGGCCCAATAGAACAGATGCACTCGCTG 60  
Db |||||  
QY 61 TGATTTTAAGACAAGCTGTATAACAGAACTCCACTGCAAGAGGNGGGCCGGGCCAGGA 120  
Db |||||  
QY 1127 TGATTTTAAGACAAGCTGTATAACAGAACTCCACTGCAAGAGGNGGGCCGGGCCAGGA 1186  
QY 121 GAATCTCCGCTTGTCCAAGACAGGGGCTTAAGAGGGTCTCCACACTGCTGTAGGGGCT 180  
Db |||||  
QY 1187 GAATCTCCGCTTGTCCAAGACAGGGGCTTAAGAGGGTCTCCACACTGCTGTAGGGGCT 1246  
QY 181 GTTGCATTTTATTTAGTAGAAGTGAAGAGGCTCTTCTCAACTTTTTCCTTGGGC 240  
Db |||||  
QY 1247 GTTGCATTTTATTTAGTAGAAGTGAAGAGGCTCTTCTCAACTTTTTCCTTGGGC 1306  
QY 241 TGGAGAAATTAGAAATCAGAAATTTCTGGAGTTTTCAGGCTATCATATATACGTATCCT 300  
Db |||||  
QY 1307 TGGAGAAATTAGAAATCAGAAATTTCTGGAGTTTTCAGGCTATCATATATACGTATCCT 1366  
QY 301 GAAAGCAACATAATTTCTTCCCTCCCTTTTAAATTTTGTGTTTCTTTTTCAGCA 360  
Db |||||  
QY 1367 GAAAGCAACATAATTTCTTCCCTCCCTTTTAAATTTTGTGTTTCTTTTTCAGCA 1426  
QY 361 TTAATCACTAAGGGCTTCAATTTAGTCCAGATTTTGTCTGGCTGCACTTAATATG 420  
Db |||||  
QY 1427 TTAATCACTAAGGGCTTCAATTTAGTCCAGATTTTGTCTGGCTGCACTTAATATG 1486  
QY 421 CCTCGCTTATTTAGCCGAGATCTGTCTTTTNTGTTTNTTTTNTTTTTCGCTCCC 480  
Db |||||  
QY 1487 CCTCGCTTATTTAGCCGAGATCTGTCTTTTNTGTTTNTTTTNTTTTTCGCTCCC 1544  
QY 481 CAAAGCTTTATCTGTCTGATCTTTTAAAGTTTGGGGCAGATTTCTGAATTTGGGCTA 540  
Db |||||  
QY 1545 CAAAGCTTTATCTGTCTGATCTTTTAAAGTTTGGGGCAGATTTCTGAATTTGGCTA 1603  
QY 541 AAAGACATGATTTTAAAGTACAGGCACTTCTTATTTCTTCTTTTAAAGTACATAGC 600  
Db |||||  
QY 1604 AAAGACATGATTTTAAAGTACAGGCACTTCTTATTTCTTCTTTTAAAGTACATAGC 1661  
QY 601 ATTAATCCCAATCTTATTTAAAGTACAGGCTTGAAGGTCACACTGCAATTTAT 660  
Db |||||  
QY 1662 ATTAATCCCAATCTTATTTAAAGTACAGGCTTGAAGGTCACACTGCAATTTAT 1721  
QY 661 AGGACCTTCTGGTGTCTGTCTGATTTTGAAGTCTGACATCTTGAAGTCTTGGC 720  
Db |||||  
QY 1722 AGGACCTTCTGGTGTCTGTCTGATTTTGAAGTCTGACATCTTGGATCTTTGC 1781  
QY 721 ATCAGAGAGGTAAAGGATTTTGAATTTTCAAGAGGAGAACACAGCGCAGAAATGAAG 780  
Db |||||  
QY 1782 ATCAGAGAGGTAAAGGATTTTGAATTTTCAAGAGGAGAACACAGCGCAGAAATGAAG 1841  
QY 781 GGCCAGGCTTACTAGGCTGTCCAGTGGAGGGCTCATGGTGGGACATGGAAAGAGGC 840  
Db |||||  
QY 1842 GGCCAGGCTTACTGA - GCTGTCCAGTGGAGGGCTCATGGTGGGACATGGAAAGAGGC 1900  
Db |||||

QY 841 AGCCTAGGCCCCCTGGGAGCCCACTGAGCAAGAGGAGCTAGTGAGCCTTTTGC 900  
Db |||||  
QY 1901 AGCCTAGGCCCCCTGGGAGCCCACTGAGCAAGAGGAGCTAGTGAGCCTTTTGC 960  
Db |||||  
QY 901 AGGAAAGGCTTAAGAAAGGAAACCAATTTAAACACAAAGAAACCTGTCCAAATGC 2020  
Db |||||  
QY 1961 AGGAAAGGCTTAAGAAAGGAAACCAATTTAAACACAAAGAAACCTGTCCAAATGC 1020  
Db |||||  
QY 961 TTTGGGAACTGTGTTTATTGCTTATAATGGGTCCCAAAATGGTAAACCTAGACTTCAG 2080  
Db |||||  
QY 2021 TTTGGGAACTGTGTTTATTGCTTATAATGGGTCCCAAAATGGTAAACCTAGACTTCAG 1080  
Db |||||  
QY 1021 GAGATGAGCAGAGAGCAAGAGAAATCTGGCTGTCTTCCATTTTCTGTTATCT 2140  
Db |||||  
QY 2081 GAGATGAGCAGAGAGCAAGAGAAATCTGGCTGTCTTCCATTTTCTGTTATCT 1140  
Db |||||  
QY 1081 CAGGTGAGCTGTAGAGGGGAGACATTAGAAAAAATGAAACAAACAAATTAAT 2200  
Db |||||  
QY 2141 CAGGTGAGCTGTAGAGGGGAGACATTAGAAAAAATGAAACAAACAAATTAAT 1200  
Db |||||  
QY 1141 GAGGTACGCTGAGGCTTGGGAGTCTCTTGACTCCACTACTTAATTCCTTTAGTGAGAAA 2260  
Db |||||  
QY 2201 GAGGTACGCTGAGGCTTGGGAGTCTCTTGACTCCACTACTTAATTCCTTTAGTGAGAAA 1260  
Db |||||  
QY 1201 CCTTTCAATTTTCTTTTATTAGAGGGGAGCTTACTGTGTGGCAAAATTCCTCAAT 2320  
Db |||||  
QY 2261 CCTTTCAATTTTCTTTTATTAGAGGGGAGCTTACTGTGTGGCAAAATTCCTCAAT 1320  
Db |||||  
QY 1261 AAGTTAATGAAAGTTGGCCAAATTTTCAAGGGGAGCTTACTGTGTGGGCTCCACATGCA 2380  
Db |||||  
QY 2321 AAGTTAATGAAAGTTGGCCAAATTTTCAAGGGGAGCTTACTGTGTGGGCTCCACATGCA 1380  
Db |||||  
QY 1321 ATGTTCAATGCGCAGCTGTCTGACACCGAGGAGTACTAGCCAGCAACAAAGGAGG 2440  
Db |||||  
QY 2381 ATGTTCAATGCGCAGCTGTCTGACACCGAGGAGTACTAGCCAGCAACAAAGGAGG 1440  
Db |||||  
QY 1381 TAGCTGAATTTCTTCTCTCTTTTACATTTCTTTTAAATAAGCAATTTAGTGTCTCAGTC 2500  
Db |||||  
QY 2441 TAGCTGAATTTCTTCTCTCTTTTACATTTCTTTTAAATAAGCAATTTAGTGTCTCAGTC 1500  
Db |||||  
QY 1441 CCTACTGAGTACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2560  
Db |||||  
QY 2501 CCTACTGAGTACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560  
Db |||||  
QY 1501 AAGGATTAACATTTTCACTGTGATGTATTTGTGTTGAGNAGAAAGAAAGTGTCTT 2619  
Db |||||  
QY 2561 AAGGATTAACATTTTCACTGTGATGTATTTGTGTTGCA - AAAAAAAGGAGTGTCTT 1620  
Db |||||  
QY 1561 TGTTTAAATTAATTTGGTTTGTGAATCCATCTTGTCTTTTCCCAATGGAACTAGTCAAT 2679  
Db |||||  
QY 2620 TGTTTAAATTAATTTGGTTTGTGAATCCATCTTGTCTTTTCCCAATGGAACTAGTCAAT 1680  
Db |||||  
QY 1621 AACCCATCTGAACTGTGAGAAACATCTGAAAGCTAGTCTATCAGCATCTGACAGG 2739  
Db |||||  
QY 2680 AACCCATCTGAACTGTGAGAAACATCTGAAAGCTAGTCTATCAGCATCTGACAGG 1740  
Db |||||  
QY 1681 TGAATTTGATGTTTCTCAGAACCAATTTCAACCCAGACAGCTGTTCTCTCTCTCTCTCTCT 2799  
Db |||||  
QY 2740 TGAATTTGATGTTTCTCAGAACCAATTTCAACCCAGACAGCTGTTCTATCTCTCTCTCTCT 1800  
Db |||||  
QY 1741 AATTAGTTGGGTTCTCTACATGCAACAAACCTGTCTCCATCTGTGACATAAGATC 2859  
Db |||||  
QY 2800 AATTAGTTGGGTTCTCTACATGCAACAAACCTGTCTCCATCTGTGACATAAGATC 1860  
Db |||||  
QY 1801 TGTGACTTGAAGTTAGTACAGCCCAACCACTTTATTTTCTATGTGTTTGTGCA 2919  
Db |||||  
QY 2860 TGTGACTTGAAGTTAGTACAGCCCAACCACTTTATTTTCTATGTGTTTGTGCA 1916  
Db |||||  
QY 1861 ACATATGATGTTTGAATAAAGTACCACTGTCTTTTAAATAAATAAATAAATAA 2975  
Db |||||  
QY 2920 ACATATGATGTTTGAATAAAGTACCACTGTCTTTTAAATAAATAAATAAATAA







```
QY 1441 CCTACTGAGTACTTTCTCTCCCTCCCTCTGATTAATCTTTCAACTTGAATTTGC 1500
Db 2501 CCTACTGAGTACTTTCTCTCCCTCCCTCTGATTAATCTTTCAACTTGAATTTGC 2560
QY 1501 AAGGATTACACATTTCACTGTGATGATATTTGTTGTCAGNAGAAAAGAAAAGTGTCTT 1560
Db 2561 AAGGATTACACATTTCACTGTGATGATATTTGTTGTCGA-AAAAAAGAAAAGTGTCTT 2619
QY 1561 TGTATAAATTAATCTGTTGTTGTAATCCATCTTTGCTTTTCCCACTGGAACCTAGTCAAT 1620
Db 2620 TGTATAAATTAATCTGTTGTTGTAATCCATCTTTGCTTTTCCCACTGGAACCTAGTCAAT 2679
QY 1621 AACCCATCTCTGAACCTGTGAGAAAACATCTGAAGAGCTAGTCTATCAGCATCTGCAGG 1680
Db 2680 AACCCATCTCTGAACCTGTGAGAAAACATCTGAAGAGCTAGTCTATCAGCATCTGCAGG 2739
QY 1681 TGAATTTGGATGTTCTCAGAACCATTTACCCAGACAGCTGTTCTATCTCTGTTTAATA 1740
Db 2740 TGAATTTGGATGTTCTCAGAACCATTTACCCAGACAGCTGTTCTATCTCTGTTTAATA 2799
QY 1741 AATTAGTTGGGTTCTCTACATGATCAACAAACCTGCTCCAACTCTGTCAACATAAAAGTC 1800
Db 2800 AATTAGTTGGGTTCTCTACATGATCAACAAACCTGCTCCAACTCTGTCAACATAAAAGTC 2859
QY 1801 TGTGACTTGAAGTTTGTAGTCAGACCCCAACCAACTTTATTTTCTATGTTTGTGCA 1860
Db 2860 TGTGACTTGAAGTTTGTAGTCAGACCCCAACCAACTTTATTTTCTATGTTTGTGCA 2919
QY 1861 ACATATGAGTGTGTTTGAATAAAGTACCAGTCTTTTATTAATAAANAARAAAAA 1916
Db 2920 ACATATGAGTGTGTTTGAATAAAGTACCAGTCTTTTATTAATAAANAARAAAAA 2975

RESULT 6
AX140845
LOCUS AX140845 2984 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 335 from Patent WO0134802.
ACCESSION AX140845
VERSION AX140845.1 GI:14280952
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
Skelky,Y.A. and Wang,A.
TITILE Compositions and methods for the therapy and diagnosis of prostate
cancer
JOURNAL Patent: WO 0134802-A 335 17-MAY-2001;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..2984
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 837 a 659 c 645 g 842 t 1 others
ORIGIN
Query Match 94.9%; Score 1836.4; DB 6; Length 2984;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1897; Conservative 0; Mismatches 12; Indels 7; Gaps 5;
QY 1 AATGGTATGCCCACTTAAGTATTACAGGTGGCCCAATAGAACAGATGCACTCGCTG 60
Db 1067 AATGGTATGCCCACTTAAGTATTACAGGTGGCCCAATAGAACAGATGCACTCGCTG 1126
QY 61 TGATTTTAAGACAAGCTGTATATAACAGAACTCCACTGCAAGGNGGCCGGCCAGGA 120
Db 1127 TGATTTTAAGACAAGCTGTATATAACAGAACTCCACTGCAAGAGGGGGCCGGCCAGGA 1186
```

```
QY 121 GAATCTCGCTTGTCCAAGACAGAGGGGCTTAAGAGGGTCTCCACACTGCTGCTAGGGCT 180
Db 1187 GAATCTCGCTTGTCCAAGACAGAGGGGCTTAAGAGGGTCTCCACACTGCTGCTAGGGCT 1246
QY 181 GTTGCAATTTTTTATTTAGTAGAAAGTGGAAAGGCTCTTCTCAACTTTTTTCCCTTGGGC 240
Db 1247 GTTGCAATTTTTTATTTAGTAGAAAGTGGAAAGGCTCTTCTCAACTTTTTTCCCTTGGGC 1306
QY 241 TGGAGAAATTTAGAATCAGAAAGTTTCTCGAGTCTTTCAGGCTATCATATATATCTGATCT 300
Db 1307 TGGAGAAATTTAGAATCAGAAAGTTTCTCGAGTCTTTCAGGCTATCATATATATCTGATCT 1366
QY 301 GAAAGCAACATAATCTTCTCTCCCTCTCTTAAAAATTTTGTGTCTCTTTTGGAGCAA 360
Db 1367 GAAAGCAACATAATCTTCTCTCCCTCTCTTAAAAATTTTGTGTCTCTTTTGGAGCAA 1426
QY 361 TTACTCACTAAAGGCTTCATTTTAGTCCAGATTTTTTAGTCTGGCTGCACTTAACCTTATG 420
Db 1427 TTACTCACTAAAGGCTTCATTTTAGTCCAGATTTTTTAGTCTGGCTGCACTTAACCTTATG 1486
QY 421 CCTCGCTTATTTAGCCCGAGATCTGCTCTTTTNTNTNTNTNTNTNTNTNTNTNTNTNTNT 480
Db 1487 CCTCGCTTATTTAGCCCGAGATCTGCTCTTTTNTNTNTNTNTNTNTNTNTNTNTNTNTNT 1544
QY 481 CAAAGCTTTATCTGTCTTGAATTTTTTAAAAAGTTTGGGGCAGATCTCGAATTTGGGCTA 540
Db 1545 CAAAGCTTTATCTGTCTTGAATTTTTTAAAAAGTTTGGGGCAGATCTCGAATTTGGGCTA 1603
QY 541 AAAGACATGCAATTTTAAAACTAGCACTCTTATTTCTTTTCTTTTAAAAATACATAGC 600
Db 1604 AAAGACATGCAATTTTAAAACTAG--CAACTCTTATTTCTTTCTTTAAAAATACATAGC 1661
QY 601 ATTAATCCCAATCCTATTTAAAGACCTTGACAGCTTGAGAAAGTCACTACTGCAATTTAT 660
Db 1662 ATTAATCCCAATCCTATTTAAAGACCTTGACAGCTTGAGAAAGTCACTACTGCAATTTAT 1721
QY 661 AGGACCTTCTGCTGTTCTGCTGTTAGTGAAGTCTGCAATCTGAGAAATCTTTGCTG 720
Db 1722 AGGACCTTCTGCTGTTCTGCTGTTAGTGAAGTCTGCAATCTGAGAAATCTTTGCTG 1781
QY 721 ATGACAGAGGTAAGAGTATTGCAATTTTACAGAGGAAGAACACAGCGCAGAAATGAAG 780
Db 1782 ATGACAGAGGTAAGAGTATTGCAATTTTACAGAGGAAGAACACAGCGCAGAAATGAAG 1841
QY 781 GGCCAGGCTTACTGAGGCTGTCAGTGGAGGCTCATGGTGGGACATGGGAAAAGAAAGC 840
Db 1842 GGCCAGGCTTACTGA-GCTGTCCAGTGGAGGCTCATGGTGGGACATGGGAAAAGAAAGC 1900
QY 841 AGCCTAGGCCCTGGGAGCCCACTGAGCAAGCAAGGAGCTGAGTGAGCCTTTTTCG 900
Db 1901 AGCCTAGGCCCTGGGAGCCCACTGAGCAAGCAAGGAGCTGAGTGAGCCTTTTTCG 1960
QY 901 AGGAAAGGCTAAGAAAAGGAAACCAATTTAAACACACAAAGAACTGTCAAATGC 960
Db 1961 AGGAAAGGCTAAGAAAAGGAAACCAATTTAAACACACAAAGAACTGTCAAATGC 2020
QY 961 TTTGGGAAGTGTGTTTATTTGCTTATAATGGGTCCCAAAATGGGTAACTAGACTTCAGA 1020
Db 2021 TTTGGGAAGTGTGTTTATTTGCTTATAATGGGTCCCAAAATGGGTAACTAGACTTCAGA 2080
QY 1021 GAGAAATGACGAGAGCAAGGAGAAATCTGGCTGCTCTCTCATTTTTCATTTCTGTTATCT 1080
Db 2081 GAGAAATGACGAGAGCAAGGAGAAATCTGGCTGCTCTCTCATTTTTCATTTCTGTTATCT 2140
QY 1081 CAGGTGAGCTGGTAGAGGGGAGACATTAGAAAATAATGAACAAACAAACAACTTACTAAT 1140
Db 2141 CAGGTGAGCTGGTAGAGGGGAGACATTAGAAAATAATGAACAAACAAACAACTTACTAAT 2200
QY 1141 GAGGTACCTGAGGCTGGGAGTCTTTGACTTCCACTACTTAATTCGGTTTGTAGTGAGAAA 1200
Db 2201 GAGGTACCTGAGGCTGGGAGTCTTTGACTTCCACTACTTAATTCGGTTTGTAGTGAGAAA 2260
QY 1201 CCTTTCAATTTTCTTTTATTAGAGGGCCAGCTTACTGTTGGTGCAAAATTTGCCAAT 1260
```

Db 2261 CTTTCAATTTCTTTTATTAGAGGCGCAGCTTACTGTGTGGGCAAAATGGCCAACAT 2320  
Qy 1261 AAGTTAATAGAAAGTTGGCCAAATTCACCCCAATTTCTGTGGTTGGGCTCACATTTGCA 1320  
Db 2321 AAGTTAATAGAAAGTTGGCCAAATTCACCCCAATTTCTGTGGTTGGGCTCACATTTGCA 2380  
Qy 1321 ATGTTCAATGCCACGTTGCTGTGACACCGAGGAGTACTAGCCAGCACAAGAGGCGAGG 1380  
Db 2381 ATGTTCAATGCCACGTTGCTGTGACACCGAGGAGTACTAGCCAGCACAAGAGGCGAGG 2440  
Qy 1381 TAGCCTGAATTTGCTTTCTGCTCTTTTACATTTCTTTTAAATAAGCAATTTAGTGTCTCAGTC 1440  
Db 2441 TAGCCTGAATTTGCTTTCTGCTCTTTTACATTTCTTTTAAATAAGCAATTTAGTGTCTCAGTC 2500  
Qy 1441 CCTACTGAGTACTTTCTCTCCCTCTCTGATTTTAAATTTCTTTTCACTTGAATTTTGC 1500  
Db 2501 CCTACTGAGTACTTTCTCTCCCTCTCTGATTTTAAATTTCTTTTCAACTTGAATTTTGC 2560  
Qy 1501 AAGGATTACACATTTCACTGTGATGATATTGTGTTCAGNGAGAAAGAAAGAGTGTCTT 1560  
Db 2561 AAGGATTACACATTTCACTGTGATGATATTGTGTTCAGNGAGAAAGAAAGAGTGTCTT 2619  
Qy 1561 TGTTTAAATTTACTTGGTTGTGAATCCATTTCTGCTTTTCCCATTTGGAAGTGTCTT 1620  
Db 2620 TGTTTAAATTTACTTGGTTGTGAATCCATTTCTGCTTTTCCCATTTGGAAGTGTCTT 2679  
Qy 1621 AACCCATCTCTGAGTGTGAGAAACATCTGAAGAGTGTCTATCAGGATCTGACAGG 1680  
Db 2680 AACCCATCTCTGAGTGTGAGAAACATCTGAAGAGTGTCTATCAGGATCTGACAGG 2739  
Qy 1681 TGAATTTGATGTTTCTCAGAACCATTTTACCAGACAGCTGTTTCTATCTGTTTAAATA 1740  
Db 2740 TGAATTTGATGTTTCTCAGAACCATTTTACCAGACAGCTGTTTCTATCTGTTTAAATA 2799  
Qy 1741 AATTAGTTGGGTTCTCTACATGATACATAACAAACCTGCTTCAATCTGTACATAAAAGTC 1800  
Db 2800 AATTAGTTGGGTTCTCTACATGATACATAACAAACCTGCTTCAATCTGTACATAAAAGTC 2859  
Qy 1801 TGTGACTTGAAGTTTGTAGTGTGAGACCCGACCAACATTTATTTCTATGTTTGTGCA 1860  
Db 2860 TGTGACTTGAAGTTTGTAGTGTGAGACCCGACCAACATTTATTTCTATGTTTGTGCA 2919  
Qy 1861 ACATATGAGTGTGTTTGAATAAAGTACCCATGCTTTTATTAATAAATAAATAAATAA 1916  
Db 2920 ACATATGAGTGTGTTTGAATAAAGTACCCATGCTTTTATTAATAAATAAATAAATAA 2975

RESULT 7  
AX200705 2984 bp DNA linear PAT 29-AUG-2001  
LOCUS  
DEFINITION Sequence 335 from Patent WO0151633.  
ACCESSION AX200705  
VERSION AX200705.1 GI:15390594  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Xu, J., Dillon, D. C., Mitcham, J. L., Harlocker, S. L., Jiang, Y., Reed, S. G., Kalos, M. D., Fanger, G. R., Day, C. H., Retter, M. W., Stolk, J. A., Skeiky, Y. A., Wang, A. and Meagher, M. J.  
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer  
JOURNAL Patent: WO 0151633-A 335 19-JUL-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1. .2984  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 837 a 659 c 645 g 842 t 1 others

ORIGIN  
Query Match 94.9%; Score 1836.4; DB 6; Length 2984;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1897; Conservative 0; Mismatches 12; Indels 7; Gaps 5;  
Qy 1 AATGTTATGCCAACTTAAGTATTATTTACAGGGTGGCCCAATAGAAACAAGATGCATCGCTG 60  
Db 1067 AATGTTATGCCAACTTAAGTATTATTTACAGGGTGGCCCAATAGAAACAAGATGCATCGCTG 1126  
Qy 61 TGATTTTAAAGCAAGCTGTATATAACAGAACTTCCACTGCAAGAGGNGGGCCGGCCAGCA 120  
Db 1127 TGATTTTAAAGCAAGCTGTATATAACAGAACTTCCACTGCAAGAGGNGGGCCGGCCAGCA 1186  
Qy 121 GAATCTCCGCTTGTCCAAAGACAGGGGCTTAAGAGGGGTCTCCACACTGCTGCTAGGGGCT 180  
Db 1187 GAATCTCCGCTTGTCCAAAGACAGGGGCTTAAGAGGGGTCTCCACACTGCTGCTAGGGGCT 1246  
Qy 181 GTTGCAATTTTATTTAGTAGAAGTGAAAGGCGCTTCTCTCAACTTTTTTCCCTTGGGC 240  
Db 1247 GTTGCAATTTTATTTAGTAGAAGTGAAAGGCGCTTCTCTCAACTTTTTTCCCTTGGGC 1306  
Qy 241 TGGAGAAATTAGATCAGAAAGTTTCTGAGTTTTCAGGCTATCATATATATCTATCTCT 300  
Db 1307 TGGAGAAATTAGATCAGAAAGTTTCTGAGTTTTCAGGCTATCATATATATCTATCTCT 1366  
Qy 301 GAAAGGCAACATAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 360  
Db 1367 GAAAGGCAACATAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1426  
Qy 361 TTACTCAGTAAAGGCGCTTCAATTTAGTCCAGATTTTATGCTGTGGCTGCACTTAATATG 420  
Db 1427 TTACTCAGTAAAGGCGCTTCAATTTAGTCCAGATTTTATGCTGTGGCTGCACTTAATATG 1486  
Qy 421 CCTCGCTTATTTAGCCGAGATCTGGTCTTTTNTGNTGNTTNTTNTTNTTNTTNTTNTTNT 480  
Db 1487 CCTCGCTTATTTAGCCGAGATCTGGTCTTTTNTGNTGNTTNTTNTTNTTNTTNTTNTT 1544  
Qy 481 CAAAGCTTATCTGCTTGTGACTTTTAAAGGTTGGGGGAGAGTCTGAAATTTGGGCTA 540  
Db 1545 CAAAGCTTATCTGCTTGTGACTTTTAAAGGTTGGGGGAGAGTCTGAAATTTGGGCTA 1603  
Qy 541 AAAGCATGCAATTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 600  
Db 1604 AAAGCATGCAATTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1661  
Qy 601 ATTAATCCAAATCTTATTTAAAGACCTGACAGCTTGAGAGGTCACCTACTGCAATTTAT 660  
Db 1662 ATTAATCCAAATCTTATTTAAAGACCTGACAGCTTGAGAGGTCACCTACTGCAATTTAT 1721  
Qy 661 AGGACCTTCTGCTGTTCTGCTGTTTACAGTCTGACAAATCTTCTGAGAAATCTTTTGC 720  
Db 1722 AGGACCTTCTGCTGTTCTGCTGTTTACAGTCTGACAAATCTTCTGAGAAATCTTTTGC 1781  
Qy 721 ATGCAGAGAGGTAAGAGTATTTGATTTTCAAGAGGAGAAACACAGCGCAGAAATGAAG 780  
Db 1782 ATGCAGAGAGGTAAGAGTATTTGATTTTCAAGAGGAGAAACACAGCGCAGAAATGAAG 1841  
Qy 781 GGCAGAGCTTACTAGGCTTCTCAGTGGAGGCTCATGGTGGGACATGGAAAGAGGCG 840  
Db 1842 GGCAGAGCTTACTGA-GCTGTCAGTGGAGGCTCATGGTGGGACATGGAAAGAGGCG 1900  
Qy 841 AGCCTTAGGCGCTTGGGAGCCAGTCCACTGACAGAGGAGGACCTGAGTGAGGCTTTTGC 900  
Db 1901 AGCCTTAGGCGCTTGGGAGCCAGTCCACTGACAGAGGAGGACCTGAGTGAGGCTTTTGC 1960  
Qy 901 AGGAAAGGCTTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 960  
Db 1961 AGGAAAGGCTTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2020  
Qy 961 TTTGGAACTGTGTTTATTTGCTTATTAAGGTTCCCAAAATGGTAACTAGACTTTTCA 1020  
Db 2021 TTTGGAACTGTGTTTATTTGCTTATTAAGGTTCCCAAAATGGTAACTAGACTTTTCA 2080

QY	1021	GAGAAATGAGCAGAGACAAAGGAGAAATCTGGCTGCTCTTCCATTTTCATTTCTGTTATCT	1080
Db	2081	GAGAAATGAGCAGAGACAAAGGAGAAATCTGGCTGCTCTTCCATTTTCATTTCTGTTATCT	2140
QY	1081	CAGGTGAGCTGGTAGAGGGGAGCATTTAGAAAAAATGAACAAACAAACAAATTTACTAAT	1140
Db	2141	CAGGTGAGCTGGTAGAGGGGAGCATTTAGAAAAAATGAACAAACAAACAAATTTACTAAT	2200
QY	1141	GAGTACGCTGAGCGCTGGGAGTCTCTTGACTCCACTACTTAATTCGTTTAGTGAGAAA	1200
Db	2201	GAGTACGCTGAGCGCTGGGAGTCTCTTGACTCCACTACTTAATTCGTTTAGTGAGAAA	2260
QY	1201	CTTTTCAATTTCTTTTATTTAGAGGGCCAGCTTACTGTTGGTGCAAAATTTGCCAACAT	1260
Db	2261	CTTTTCAATTTCTTTTATTTAGAGGGCCAGCTTACTGTTGGTGCAAAATTTGCCAACAT	2320
QY	1261	AGTTAATAGAAATTTGGCCAAATTTCAACCCATTTTCTGTGTTTGGGCTCCACATTGCA	1320
Db	2321	AGTTAATAGAAATTTGGCCAAATTTCAACCCATTTTCTGTGTTTGGGCTCCACATTGCA	2380
QY	1321	ATGTTCAATGCCAGCTGCTGCACACCGAGTACTAGCCAGCACAACAAAGGCAGGG	1380
Db	2381	ATGTTCAATGCCAGCTGCTGCACACCGAGTACTAGCCAGCACAACAAAGGCAGGG	2440
QY	1381	TAGCCTGAATTTGCTTCTGCTCTTTTACATTTCTTTTAAATAGCAATTTAGTGTCTCAGTC	1440
Db	2441	TAGCCTGAATTTGCTTCTGCTCTTTTACATTTCTTTTAAATAGCAATTTAGTGTCTCAGTC	2500
QY	1441	CCTACTGAGTACTCTTTCTCTCCCTCTCTGAAATTTCTTCAACTTGAATTTGCAATTTGC	1500
Db	2501	CCTACTGAGTACTCTTTCTCTCCCTCTCTGAAATTTCTTCAACTTGAATTTGCAATTTGC	2560
QY	1501	AAGGATTACACATTTCACTGTGATGATATTTGTTGCGAGGAGAAAGAAAGTGTCTT	1560
Db	2561	AAGGATTACACATTTCACTGTGATGATATTTGTTGCGA-AAAAAAGAAAGTGTCTT	2619
QY	1561	TGTTTAAATTTACTTGGTTTGTGAATCCATTTCTGCTTTTCCCACTGGAATAGTCAAT	1620
Db	2620	TGTTTAAATTTACTTGGTTTGTGAATCCATTTCTGCTTTTCCCACTGGAATAGTCAAT	2679
QY	1621	AACCCATCTGAACTGGTAGAAAAACATCTGAAGAGTACTATCAGCATCTGACAGG	1680
Db	2680	AACCCATCTGAACTGGTAGAAAAACATCTGAAGAGTACTATCAGCATCTGACAGG	2739
QY	1681	TGAATTTGATGTTCTCAGAACATTTTCCAGAGAGCTGTTCTATCTCTGTTTAAATA	1740
Db	2740	TGAATTTGATGTTCTCAGAACATTTTCCAGAGAGCTGTTCTATCTCTGTTTAAATA	2799
QY	1741	AATTAGTTTGGGTTCTCTACATGCATAACAAACCTGCTCCAATCTGTGCACATAAAGTC	1800
Db	2800	AATTAGTTTGGGTTCTCTACATGCATAACAAACCTGCTCCAATCTGTGCACATAAAGTC	2859
QY	1801	TGTGACTTGAAGTTTGTAGTACGACCCCAACCAATTTATTTTCTATGTTTGTGCA	1860
Db	2860	TGTGACTTGAAGTTTGTAGTACGACCCCAACCAATTTATTTTCTATGTTTGTGCA	2919
QY	1861	ACATATGAGTGTGTTTGAATAAAGTACCATGCTTTTATTTAAAAAAGAAAAA	1916
Db	2920	ACATATGAGTGTGTTTGAATAAAGTACCATGCTTTTATTTAAAAAAGAAAAA	2975
RESULT 8			
AX267361			
LOCUS	AX267361	2984 bp	DNA
DEFINITION	Sequence 335 from Patent WO0173032.	linear	PAT 26-OCT-2001
ACCESSION	AX267361		
VERSION	AX267361.1	GI:16516135	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		

REFERENCE	1										
AUTHORS	Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y., Kalos, M.D., Fanger, G.R., Retter, M.W., Stolk, J.A., Day, C.H., Vedvick, T.S., Carter, D., Li, S.X., Wang, A., Skeiky, Y.A., Hepler, W.T. and Henderson, R.A.										
TITLE	Compositions and methods for the therapy and diagnosis of prostate cancer										
JOURNAL	Patent: WO 0173032-A 335 04-OCT-2001; CORIXA CORPORATION (US)										
FEATURES	Location/Qualifiers										
source	1..2984										
	/organism="Homo sapiens"										
	/mol_type="genomic DNA"										
	/db_xref="taxon:9606"										
BASE COUNT	837 a	659 c	842 t	1 others							
ORIGIN											
Query Match	94.9%; Score 1836.4; DB 6; Length 2984;										
Best Local Similarity	99.0%; Pred. No. 0;										
Matches 1897;	Conservative	0;	Mismatches	12;	Indels	7;	Gaps	5;			
QY	1	AATGGTATGCCAACTTAAAGTATTACAGGGTGGCCCAATAAGAACAAAGATGCACTCGCTG	60								
Db	1067	AATGGTATGCCAACTTAAAGTATTACAGGGTGGCCCAATAAGAACAAAGATGCACTCGCTG	1126								
QY	61	TGATTTTAAAGCAAGCTGTATAACAGAACTCCACTGCAAGAGGNGGCGCGGCCAGGA	120								
Db	1127	TGATTTTAAAGCAAGCTGTATAACAGAACTCCACTGCAAGAGGNGGCGCGGCCAGGA	1186								
QY	121	GAATCTCCGCTTGTCCAAACAGAGGGCCCTAAGAGAGGCTCTCCACACTGCTGCTAGGGGCT	180								
Db	1187	GAATCTCCGCTTGTCCAAACAGAGGGCCCTAAGAGAGGCTCTCCACACTGCTGCTAGGGGCT	1246								
QY	181	GTTGCAATTTTTTATTAGTAGAAGTGGAAAGCGCTCTTCTCAACTTTTTTCCCTTGGGC	240								
Db	1247	GTTGCAATTTTTTATTAGTAGAAGTGGAAAGCGCTCTTCTCAACTTTTTTCCCTTGGGC	1306								
QY	241	TGGAGAAATTTAGAAATCAGAAAGTTTCTCGAGTTTTCAGGCTATCATATACCTGATCCT	300								
Db	1307	TGGAGAAATTTAGAAATCAGAAAGTTTCTCGAGTTTTCAGGCTATCATATACCTGATCCT	1366								
QY	301	GAAAGGCAACATAATCTTCTTCCCTCTCTTTTAAAAATTTGTGTCTCTTTTTCAGCAAA	360								
Db	1367	GAAAGGCAACATAATCTTCTTCCCTCTCTTTTAAAAATTTGTGTCTCTTTTTCAGCAAA	1426								
QY	361	TTACTCATAAAGGGCTTCAATTTAGTCCAGATTTTATGCTGGCGCACTTAACCTTAGT	420								
Db	1427	TTACTCATAAAGGGCTTCAATTTAGTCCAGATTTTATGCTGGCGCACTTAACCTTAGT	1486								
QY	421	CCTCGCTTATTTAGCCCGAGATCTGCTCTTTTNTGTGTTTTTTTTTTTTCGCTCTCCC	480								
Db	1487	CCTCGCTTATTTAGCCCGAGATCTGCTCTTTTNTGTGTTTTTTTTTTTTCGCTCTCCC	1544								
QY	481	CAAAGCTTTATCTGCTTGACATTTTAAAAAAAGTTTGGGGGAGATTTCTGAATTTGGGCTA	540								
Db	1545	CAAAGCTTTATCTGCTTGACATTTTAAAAAAAGTTTGGGGGAGATTTCTGAATTTGGGCTA	1603								
QY	541	AAAGACATGCATTTTAAAGCTAGGCAACTCTTATTTCTTTTCTTTAAAAATACATAGC	600								
Db	1604	AAAGACATGCATTTTAAAGCTAGGCAACTCTTATTTCTTTTCTTTAAAAATACATAGC	1661								
QY	601	ATTAAATCCCAATCTTATTTAAAGACCTGACAGCTTGAGAGGTCACCTACTGCAATTTAT	660								
Db	1662	ATTAAATCCCAATCTTATTTAAAGACCTGACAGCTTGAGAGGTCACCTACTGCAATTTAT	1721								
QY	661	AGGACCTTCTGCTGGTTCCTGCTGTTTACGTTTGAAGTCTGCACAACTCTTGAGAATCTTTGC	720								
Db	1722	AGGACCTTCTGCTGGTTCCTGCTGTTTACGTTTGAAGTCTGCACAACTCTTGAGAATCTTTGC	1781								
QY	721	ATGCAGAGAGGTAAAGGTTATGGATTTTTCACAGGGAAGAACACAGCGCAGATGAAG	780								
Db	1782	ATGCAGAGAGGTAAAGGTTATGGATTTTTCACAGGGAAGAACACAGCGCAGATGAAG	1841								

QY 781 GGCCAGGCTTACTGAGCTGTCCAGTGGAGGGCTCATGGTGGGACATGGAAGAAGGC 840  
Db 1842 GGCCAGGCTTACTGA - CTTGTCCAGTGGAGGGCTCATGGTGGGACATGGAAGAAGGC 1900  
QY 841 AGCTAGGCTTGGGAGCCAGTCCACTGAGCAAGCAAGGACTGAGTGAGCTTTTGC 900  
Db 1901 AGCTAGGCTTGGGAGCCAGTCCACTGAGCAAGCAAGGACTGAGTGAGCTTTTGC 1960  
QY 901 AGGAAAAGGCTTAAGAAAAGGAAAACCATCTTAACACAAAGAACTGTCCAAATGC 960  
Db 1961 AGGAAAAGGCTTAAGAAAAGGAAAACCATCTTAACACAAAGAACTGTCCAAATGC 2020  
QY 961 TTTGGGAAGTGTGTTTATTCCTATATATGGTCCCCAAAATGGTAACTAGACTTCAGA 1020  
Db 2021 TTTGGGAAGTGTGTTTATTCCTATATATGGTCCCCAAAATGGTAACTAGACTTCAGA 2080  
QY 1021 GAGNATGAGCAGAGCAAGGAGAAATCTGGCTGCTCCATTTTCATCTGTTATCT 1080  
Db 2081 GAGNATGAGCAGAGCAAGGAGAAATCTGGCTGCTCCATTTTCATCTGTTATCT 2140  
QY 1081 CAGGTGAGCTGTGAGGGGAGACATTAAGAAAATAATGAACAAACAAATTAAT 1140  
Db 2141 CAGGTGAGCTGTGAGGGGAGACATTAAGAAAATAATGAACAAACAAATTAAT 2200  
QY 1141 GAGGTACGCTGAGGCTGGAGTCTCTTGACTCCACTTAATTCGGTTTAGTGAGAA 1200  
Db 2201 GAGGTACGCTGAGGCTGGAGTCTCTTGACTCCACTTAATTCGGTTTAGTGAGAA 2260  
QY 1201 CTTTCAATTTCTTTTATAGAGGCGAGCTACTGTTGGTGGCAAAATTCGCAACAT 1260  
Db 2261 CTTTCAATTTCTTTTATAGAGGCGAGCTACTGTTGGTGGCAAAATTCGCAACAT 2320  
QY 1261 AAGTTAATAGAAAGTTGGCCAAATTCACCCCAATTTCTGTGGTTCGGCTCCACATGCA 1320  
Db 2321 AAGTTAATAGAAAGTTGGCCAAATTCACCCCAATTTCTGTGGTTCGGCTCCACATGCA 2380  
QY 1321 ATGTTCAATGCCAGTGTCTGACACCGAGTACTAGCCAGCAACAAAGGCGAGG 1380  
Db 2381 ATGTTCAATGCCAGTGTCTGACACCGAGTACTAGCCAGCAACAAAGGCGAGG 2440  
QY 1381 TAGCCTGAATGCTTCTGCTCTTTTACATTTCTTTTAAATAGCAATTTAGTCTCAGTC 1440  
Db 2441 TAGCCTGAATGCTTCTGCTCTTTTACATTTCTTTTAAATAGCAATTTAGTCTCAGTC 2500  
QY 1441 CCTACTGAGTACTTCTCTCCCTCTCTGAAATTAATTTCTTCAACTTGAATTTGC 1500  
Db 2501 CCTACTGAGTACTTCTCTCCCTCTCTGAAATTAATTTCTTCAACTTGAATTTGC 2560  
QY 1501 AAGATTACACATTTCACTGTGATATATTTGTTGCAAGGAAAAGAAAAGTGTCTT 1560  
Db 2561 AAGATTACACATTTCACTGTGATATATTTGTTGCA - AAAAAAAGAAAAGTGTCTT 2619  
QY 1561 TGTTTAAATTAATTTGTTGTAATCCATCTGCTTTTCCCATTTGCAACTAGTCAAT 1620  
Db 2620 TGTTTAAATTAATTTGTTGTAATCCATCTGCTTTTCCCATTTGCAACTAGTCAAT 2679  
QY 1621 AACCCATCTCTGAACGTGTAAGAAAACATCTGAAGAGTACTATCAGCACTGACAGG 1680  
Db 2680 AACCCATCTCTGAACGTGTAAGAAAACATCTGAAGAGTACTATCAGCACTGACAGG 2739  
QY 1681 TGAATGGATGGTTCTCAGAACCAATTTACCCAGAGCGCTGTTTCTATCTGTTTAAATA 1740  
Db 2740 TGAATGGATGGTTCTCAGAACCAATTTACCCAGAGCGCTGTTTCTATCTGTTTAAATA 2799  
QY 1741 AATTAGTTGGGTTCTTACATGATCAACAAACCTGCTCCATCTGTCACATAAAGTC 1800  
Db 2800 AATTAGTTGGGTTCTTACATGATCAACAAACCTGCTCCATCTGTCACATAAAGTC 2859  
QY 1801 TGTGACTTGAAGTTTGTAGTCAGCACCCCAACCAACTTTATTTTCTATGTTTGTGCA 1860  
Db 2860 TGTGACTTGAAGTTTGTAGTCAGCACCCCAACCAACTTTATTTTCTATGTTTGTGCA 2919  
QY 1861 ACATATGAGTGTGTTGAAAATAAAGTACCAATGCTTTTATTAATAAATAAATAA 1916

Db 2920 ACATATGAGTGTGTTGAAAATAAAGTACCAATGCTTTTATTAATAAATAAATAA 2975

RESULT 9  
AC012574/c  
LOCUS

DEFINITION Homo sapiens chromosome 8, clone RP11-213G6, linear PRI 01-APR-2002  
AC012574  
AC012574.7 GI:19852145  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 80586)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
JOURNAL Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 2 (bases 1 to 80586)  
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
TITLE Homo sapiens chromosome 8, clone RP11-213G6  
JOURNAL Unpublished

REFERENCE 3 (bases 1 to 80586)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
JOURNAL Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE  
JOURNAL

REFERENCE 3 (bases 1 to 80586)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
JOURNAL Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS

REFERENCE 3 (bases 1 to 80586)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
JOURNAL Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE  
JOURNAL

REFERENCE 3 (bases 1 to 80586)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
JOURNAL Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## COMMENT

On Apr 1, 2002 this sequence version replaced gi:15213914.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIER

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2354  
Center clone name: 213\_G\_6  
-----

Only the first 80.6 kilobases of this clone is being submitted.  
The remainder overlaps either accession number AC051642 [WICGR  
project L8828] or accession number AC012119 [WICGR project L2190].

## FEATURES

## source

1. .80586  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"

/clone="RP11-213G6"  
/clone\_lib="RPC1-11 Human Male BAC"

repeat\_region  
69. .98

repeat\_region  
/rpt\_family="TCTA)n"  
complement(366. .557)

repeat\_region  
/rpt\_family="MIR"  
623. .654

repeat\_region  
/rpt\_family="(TAAAAA)n"  
complement(1315. .1430)

repeat\_region  
/rpt\_family="MIR"  
complement(1613. .1847)

repeat\_region  
/rpt\_family="L3"  
1850. .1875

repeat\_region  
/rpt\_family="(TTC)n"  
complement(1905. .2003)

repeat\_region  
/rpt\_family="L3"  
complement(2701. .2738)

repeat\_region  
/rpt\_family="MIR"  
3467. .3594

repeat\_region  
/rpt\_family="FLAM\_A"  
4305. .4606

repeat\_region  
/rpt\_family="AluJo"  
complement(4624. .4758)

repeat\_region  
/rpt\_family="MER5B"  
4884. .5265

repeat\_region  
/rpt\_family="TBE1C"  
complement(6193. .6343)

repeat\_region  
/rpt\_family="AluJo"  
complement(6950. .7155)

repeat\_region  
/rpt\_family="AluY"  
complement(7674. .7984)

repeat\_region  
/rpt\_family="AluX"  
complement(8132. .8556)

repeat\_region  
/rpt\_family="L1PB3"  
10228. .10254

repeat\_region  
/rpt\_family="(A)n"  
12739. .12759

repeat\_region  
/rpt\_family="GC-rich"  
13208. .13284

repeat\_region  
/rpt\_family="CT-rich"  
complement(13943. .14098)

repeat\_region  
/rpt\_family="FRAM"  
14128. .14390

repeat\_region  
/rpt\_family="L1MC4"  
complement(14600. .14741)

repeat\_region  
/rpt\_family="L1MD3"  
complement(14746. .15043)

repeat\_region  
/rpt\_family="AluJb"  
complement(15044. .15527)

repeat\_region  
/rpt\_family="L1MC/D"  
15647. .15833

repeat\_region  
/rpt\_family="L1MC/D"  
15834. .16178

repeat\_region  
/rpt\_family="MER7A"  
16179. .16276

repeat\_region  
/rpt\_family="L1MC/D"  
16720. .16859

repeat\_region  
/rpt\_family="L2"  
17036. .17556

repeat\_region  
/rpt\_family="MLT1D"  
17697. .17757

repeat\_region  
complement(17882. .17964)

repeat\_region  
/rpt\_family="MIR3"  
complement(18120. .18205)

repeat\_region  
/rpt\_family="MER74A"  
complement(18206. .18505)

repeat\_region  
/rpt\_family="AluSp"  
complement(18506. .18674)

repeat\_region  
/rpt\_family="MER74A"  
complement(18966. .19348)

repeat\_region  
/rpt\_family="MLT1A1"  
19021. .19032

unsure  
/note="<30 qual SNGL region"  
complement(19528. .19617)

repeat\_region  
/rpt\_family="MIR"  
20289. .20525

repeat\_region  
/rpt\_family="AluJb"  
20526. .20545

repeat\_region  
/rpt\_family="(A)n"  
complement(20548. .20630)

repeat\_region  
/rpt\_family="MER5A"  
complement(22022. .22171)

repeat\_region  
/rpt\_family="MIR"  
22431. .22598

repeat\_region  
/rpt\_family="AluJb"  
23668. .23859

repeat\_region  
/rpt\_family="L1ME3A"  
23860. .24152

repeat\_region  
/rpt\_family="AluSp"  
24153. .24626

repeat\_region  
/rpt\_family="L1ME3A"  
complement(25149. .25467)

repeat\_region  
/rpt\_family="MLT1J"  
27162. .27232

repeat\_region  
/rpt\_family="MIR"  
27510. .27657

repeat\_region  
/rpt\_family="L1ME2"  
27658. .27967

repeat\_region  
/rpt\_family="AluJo"  
27968. .28021

repeat\_region  
/rpt\_family="L1ME2"  
complement(28389. .28516)

repeat\_region  
/rpt\_family="MIR3"  
29420. .29441

repeat\_region  
/rpt\_family="AT-rich"  
30731. .31010

repeat\_region  
/rpt\_family="AluJo"  
31022. .31046

repeat\_region  
/rpt\_family="(TAAA)n"  
33517. .33834

repeat\_region  
/rpt\_family="AluY"  
33935. .33998

repeat\_region  
/rpt\_family="GA-rich"  
complement(37955. .38196)

repeat\_region  
/rpt\_family="AluJb"  
38197. .38216

repeat\_region  
/rpt\_family="(TTTA)n"  
complement(38217. .38333)

Query Match 94.7%; Score 1834.2; DB 9; Length 80586;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1887; Conservative 0; Mismatches 13; Indels 8; Gaps 4;

QY 1 AATGTTATGCCAACTTAAGTATTTACAGGTGGCCCAATAGAACAAAGATGACACTCGCTG 60

Db 10702 AATGTTATGCCAACTTAAGTATTTACAGGTGGCCCAATAGAACAAAGATGACACTCGCTG 10643

QY 61 TGATTTTAAAGCAAGCTGTATATAACAGACTCCACTGCAAGAGGNGGCGCGGCCAGGA 120

Db 10642 TGATTTTAAAGCAAGCTGTATATAACAGACTCCACTGCAAGAGGNGGCGCGGCCAGGA 10583

QY 121 GAATCTCGCTTGTCCCAAGACAGGGGCTAAGAGGCTCTCCACACTGCTGCTAGGGCT 180  
Db 10582 GAATCTCGCTTGTCCCAAGACAGGGGCTAAGAGGCTCTCCACACTGCTGCTAGGGCT 10523  
QY 181 GTTCGATTTTATTATAGTAGAAGAGGCTCTCTCAACTTTTTCCTTGGGC 240  
Db 10522 GTTCGATTTTATTATAGTAGAAGAGGCTCTCTCAACTTTTTCCTTGGGC 10463  
QY 241 TGGAGAAATTTAGAAATCAGAGTTTCTCGAGTTTTCAGGCTATCATATATACGTATCCT 300  
Db 10462 TGGAGAAATTTAGAAATCAGAGTTTCTCGAGTTTTCAGGCTATCATATATACGTATCCT 10403  
QY 301 GAAAGGCAACATAATCTCTCCCTCCCTTTTAAATTTTGGTTCCCTTTTTCGAGCAA 360  
Db 10402 GAAAGGCAACATAATCTCTCCCTCCCTTTTAAATTTTGGTTCCCTTTTTCGAGCAA 10343  
QY 361 TTACTCACTAAGGGCTTCATTTTAGTCCAGATTTTGTAGTCTGGCTGCACCTAATTAAG 420  
Db 10342 TTACTCACTAAGGGCTTCATTTTAGTCCAGATTTTGTAGTCTGGCTGCACCTAATTAAG 10283  
QY 421 CCTCGCTATTATAGCCGAGATCTGGTC---TTTTTTTNTGTTTTTTTTTTTTTCCGTC 476  
Db 10282 CCTCGCTATTATAGCCGAGATCTGGTC---TTTTTTTNTGTTTTTTTTTTTTTCCGTC 10223  
QY 477 TCCCAAGCTTTATCTGTCTGTGATTTTAAAGTTTGGGGGAGATTTCTGAATGG 536  
Db 10222 TCCCAAGCTTTATCTGTCTGTGATTTTAAAGTTTGGGGGAGATTTCTGAAT-G 10164  
QY 537 GCTAAGAGCATGATTTTAAACTAGGCACTTCTTATTTCTTTTCTTTTAAATATACA 596  
Db 10163 GCTAAGAGCATGATTTTAAACTAG--CAACTCTTATTTCTTTTCTTTTAAATATACA 10106  
QY 597 TAGCATTAATCCCAATCTTATTAAGACCTGACAGCTTTGAGAAGTCACTACTGCAT 656  
Db 10105 TAGCATTAATCCCAATCTTATTAAGACCTGACAGCTTTGAGAAGTCACTACTGCAT 10046  
QY 657 TTATAGACCTTCTGGTGTCTGTCTTACGTTTGAAGTCTGACAACTCTGAGAATCT 716  
Db 10045 TTATAGACCTTCTGGTGTCTGTCTTACGTTTGAAGTCTGACAACTCTGAGAATCT 9986  
QY 717 TTGCATCAGAGGAGTAAGAGTATGAGTTTTCAGAGGAAGAACACAGGCGAGAT 776  
Db 9985 TTGCATCAGAGGAGTAAGAGTATGAGTTTTCAGAGGAAGAACACAGGCGAGAT 9926  
QY 777 GAAGGGCAGGCTTACTGAGGCTGTCCAGTGAGGGCTCATGGTGGGACATGGAAGA 836  
Db 9925 GAAGGGCAGGCTTACTGA-GCTGTCCAGTGAGGGCTCATGGTGGGACATGGAAGA 9867  
QY 837 AGGAGGCTAGGCTGGGAGGCCAGTCCACTGAGCAAGGAGGAGTGAAGGCTT 896  
Db 9866 AGGAGGCTAGGCTGGGAGGCCAGTCCACTGAGCAAGGAGGAGTGAAGGCTT 9807  
QY 897 TTGCAGAAAGGCTAAGAAAGGAAACCATTTCTAAGCAACAAGAACTGTCGA 956  
Db 9806 TTGCAGAAAGGCTAAGAAAGGAAACCATTTCTAAGCAACAAGAACTGTCGA 9747  
QY 957 ATGCTTTGGAACTGTGTTTATGCTTATTAATGGTCCCCAAATGGTAACTAGACTT 1016  
Db 9746 ATGCTTTGGAACTGTGTTTATGCTTATTAATGGTCCCCAAATGGTAACTAGACTT 9687  
QY 1017 CAGAGAAATCAGCAGAGCAAGAGAAATCTGGCTGCTTCCATTTTCACTCTGTT 1076  
Db 9686 CAGAGAAATCAGCAGAGCAAGAGAAATCTGGCTGCTTCCATTTTCACTCTGTT 9627  
QY 1077 ATCTCAGGCTAGGCTAGAGGGAGACATTTAGAAAAAATGAAACAAACAACATTAAC 1136  
Db 9626 ATCTCAGGCTAGGCTAGAGGGAGACATTTAGAAAAAATGAAACAAACAACATTAAC 9567  
QY 1137 TAATGAGGTACGTAGGCTGGAGTCTTGAATCTTGAATCTTAAATCCGTTTAGTGA 1196  
Db 9566 TAATGAGGTACGTAGGCTGGAGTCTTGAATCTTGAATCTTAAATCCGTTTAGTGA 9507  
QY 1197 GAACCTTCAATTTCTTTTATAGAGGGCCAGCTTACTGTGTGGCAAAATGGCA 1256

Db 9506 GAAACCTTCAATTTCTTTATTAGAGGGCCAGCTTACTTGTGGTGCAAAATGGCA 9447  
QY 1257 ACATAAGTTAATAGAAAGTTGGCCAAATTTCAACCATAATTTCTGTGTTGGGCTCCACAT 1316  
Db 9446 ACATAAGTTAATAGAAAGTTGGCCAAATTTCAACCATAATTTCTGTGTTGGGCTCCACAT 9387  
QY 1317 TCGAATGTTCAATGCCAGTGTCTGCACCGAGCGGAGTACTAGCCAGCAAAAGGC 1376  
Db 9386 TCGAATGTTCAATGCCAGTGTCTGCACCGAGCGGAGTACTAGCCAGCAAAAGGC 9327  
QY 1377 AGGAGTCCCTGAATTTCTGCTCTTTTACATTTCTTTTAAATAAGCAATTTAGTGCTC 1436  
Db 9326 AGGAGTCCCTGAATTTCTGCTCTTTTACATTTCTTTTAAATAAGCAATTTAGTGCTC 9267  
QY 1437 AGTCCCTACTAGTACTCTTTCTCTCCCTCTCTCTGAAATTTAAATTTCTTCAACTGCAAT 1496  
Db 9266 AGTCCCTACTAGTACTCTTTCTCTCCCTCTCTCTGAAATTTAAATTTCTTCAACTGCAAT 9207  
QY 1497 TTGCAAGGATTACACATTTCACTGTGATGTATTTGTTGCGAGGAAAGAAAGTG 1556  
Db 9206 TTGCAAGGATTACACATTTCACTGTGATGTATTTGTTGCGAATAAATAAAGTG 9147  
QY 1557 TCTTTGTTTAAATTTACTGTTTGTGAATCCATCTTGTCTTTTCCCAATTTGGAATAGT 1616  
Db 9146 TCTTTGTTTAAATTTACTGTTTGTGAATCCATCTTGTCTTTTCCCAATTTGGAATAGT 9087  
QY 1617 CATTAACCCATCTCGAATCTGTAGAAAAACATCTGAAAGAGCTAGTCTATCAGCATCTGA 1676  
Db 9086 CATTAACCCATCTCGAATCTGTAGAAAAACATCTGAAAGAGCTAGTCTATCAGCATCTGA 9027  
QY 1677 CAGGTGAATGGAATGTTCTCAGAACCATTTTCCACCCAGACAGCTGTTTCTATCTGTT 1736  
Db 9026 CAGGTGAATGGAATGTTCTCAGAACCATTTTCCACCCAGACAGCTGTTTCTATCTGTT 8967  
QY 1737 AATAAATTTAGTTGGTCTCTTACATGCAATCAACAACTCTGCTCAATCTGTCAATAAA 1796  
Db 8966 AATAAATTTAGTTGGTCTCTTACATGCAATCAACAACTCTGCTCAATCTGTCAATAAA 8907  
QY 1797 AGTCTGCTGCTGAGTTTGTAGTCTCAGACCCCAACCAATTTTCTATGTTGTTT 1856  
Db 8906 AGTCTGCTGCTGAGTTTGTAGTCTCAGACCCCAACCAATTTTCTATGTTGTTT 8847  
QY 1857 TCGACATATGATGTTTGAATAAATAAGTACCCATGCTTTATTTAA 1904  
Db 8846 TCGACATATGATGTTTGAATAAATAAGTACCCATGCTTTATTTAGA 8799

## RESULT 10

AC051642  
LOCUS AC051642 174445 bp DNA linear PRI 17-JAN-2002  
DEFINITION Homo sapiens chromosome 8, clone RP11-583M2, complete sequence.  
AC051642  
ACCESSION AC051642.5 GI:17975383  
VERSION AC051642.5  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 174445)  
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE Homo sapiens chromosome 8, clone RP11-583M2  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 174445)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F.,  
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,  
Campoliano, A., Castelle, A., Choe, Y., Colangelo, M., Collins, S.,  
Collins, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,

Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
 Levine, R., Liu, G., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
 Meldrim, J., Menes, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,  
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
 O'Neil, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N.,  
 Pisani, C., Pollara, V., Raymond, C., Riley, K., Rogov, P., Rothman, D.,  
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
 Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

## Direct Submission

Submitted (15-APR-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE  
AUTHORS

3 (bases 1 to 174445)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,  
 Brown, A., Camarata, J., Campoliano, A., Chang, J., Chazaro, B.,  
 Choepel, Y., Collarato, M., Collins, S., Collumore, A., Cook, A.,  
 Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,  
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hages, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
 Jones, C., Kamat, A., Karatas, A., Kells, C., Lakocque, K.,  
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,  
 Menes, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,  
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,  
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,  
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## Direct Submission

Submitted (17-JAN-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

On Dec 21, 2001 this sequence version replaced gi:16905345.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIR

Web site: <http://www.seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L8828

Center clone name: 593\_M\_2

## FEATURES

## source

Location/Qualifiers

1. .174445

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosomes="8"

/maps="8"

/clone="RP11-583M2"

/clone\_lib="RP11-11 Human Male BAC"

1560..1586

/rpt\_family="(T)n"

3258..3682

/rpt\_family="LIPB3"

3830..4140

/rpt\_family="AluSx"

4659..4864

/rpt\_family="AluY"

5471..5621

/rpt\_family="AluJo"

repeat\_region complement(6549..6930)  
 repeat\_region /rpt\_family="THE1C"  
 7056..7190  
 repeat\_region /rpt\_family="MER5B"  
 complement(7208..7509)  
 repeat\_region /rpt\_family="AluJo"  
 complement(8220..8347)  
 repeat\_region /rpt\_family="FLAM\_A"  
 9076..9113  
 repeat\_region /rpt\_family="MIR"  
 9811..9909  
 repeat\_region /rpt\_family="L3"  
 9939..9964  
 repeat\_region /rpt\_family="(TCAA)n"  
 9967..10205  
 repeat\_region /rpt\_family="L3"  
 10384..10499  
 repeat\_region /rpt\_family="MIR"  
 11160..11191  
 repeat\_region /rpt\_family="(TTTTTA)n"  
 11257..11448  
 repeat\_region /rpt\_family="MIR"  
 11716..11745  
 repeat\_region /rpt\_family="(TAGA)n"  
 12412..12602  
 repeat\_region /rpt\_family="MIR3"  
 13499..13507  
 unsure /note="probably AAATATTTT"  
 repeat\_region /rpt\_family="L3"  
 13810..15292  
 repeat\_region /rpt\_family="L3"  
 14276..14280  
 unsure /note="30 qual SNGL region"  
 repeat\_region /rpt\_family="AluJo"  
 15314..15456  
 repeat\_region /rpt\_family="L3"  
 15457..15492  
 repeat\_region /rpt\_family="L3"  
 15493..15517  
 repeat\_region /rpt\_family="(CA)n"  
 15518..16411  
 repeat\_region /rpt\_family="L3"  
 16473..16768  
 repeat\_region /rpt\_family="AluSx"  
 16775..16892  
 repeat\_region /rpt\_family="L3"  
 complement(17069..17309)  
 repeat\_region /rpt\_family="MIR"  
 19837..20309  
 repeat\_region /rpt\_family="MER21B"  
 20310..20670  
 repeat\_region /rpt\_family="MER1B"  
 20671..21014  
 repeat\_region /rpt\_family="MER21B"  
 complement(21715..22026)  
 repeat\_region /rpt\_family="AluJo"  
 22091..22127  
 repeat\_region /rpt\_family="AT rich"  
 complement(22190..22332)  
 repeat\_region /rpt\_family="MIR"  
 22455..22558  
 repeat\_region /rpt\_family="L3"  
 22801..25068  
 repeat\_region /rpt\_family="L3MB7"  
 25186..26181  
 repeat\_region /rpt\_family="L3MB1"  
 26242..26524  
 repeat\_region /rpt\_family="MERS8B"  
 28103..28399  
 repeat\_region /rpt\_family="AluSx"  
 complement(28945..28999)  
 repeat\_region /rpt\_family="MIR"  
 29087..29260  
 repeat\_region /rpt\_family="(GAAA)n"  
 complement(29655..29755)  
 repeat\_region

repeat_region	/rpt family="L2"	29959.30246	Db	1651		GCTAAAGACATGCATTTTAAACCTAG--CAACTCTTTATTTCTTTCTTTTAAATAACA	1708
repeat_region	/rpt family="AluSx"	30493.36518	Qy	597		TAGCATTAAATCCCAATCCTATTTAAAGACCTGACAGCTTGAGAGGTCACCTACTGTCAT	656
repeat_region	/rpt family="LIP2"	complement(36649.36820)	Db	1709		TAGCATTAAATCCCAATCCTATTTAAAGACCTGACAGCTTGAGAGGTCACCTACTGTCAT	1768
repeat_region	/rpt family="MIR"	36850.37060	Qy	657		TTATAGGACCTTCTGGTGGTTCTGCTGTCTGTTTAAAGCTTGACAATCTTTGAGAATCT	716
repeat_region	/rpt family="LIP13"	complement(37416.37545)	Db	1769		TTATAGGACCTTCTGGTGGTTCTGCTGTCTGTTTAAAGCTTGACAATCTTTGAGAATCT	1828
repeat_region	/rpt family="LIPB3"	complement(37553.37637)	Qy	717		TTGCATGACAGAGAGTAAAGAGTATTTGATTTTCAAGAGGAGTAAACACAGCGGAGAT	776
repeat_region	/rpt family="MIR"	38668.39223	Db	1829		TTGCATGACAGAGAGTAAAGAGTATTTGATTTTCAAGAGGAGTAAACACAGCGGAGAT	1888
repeat_region	/rpt family="L1MB3A"	39363.39803	Qy	777		GAAGGCCAGGCTTACTGAGGCTGTCCAGTGGAGGCTCATGGGTGGGACATGGAAGA	836
repeat_region	/rpt family="L1MB3A"	39807.39923	Db	1889		GAAGGCCAGGCTTACTGA-GCTGTCCAGTGGAGGCTCATGGGTGGGACATGGAAGA	1947
repeat_region	/rpt family="L1MC1"	39930.40596	Qy	837		AGGCAGCTTAGGCCCTGGGAGCCAGTCCACTGAGCAGCAAGGAGCTGAGTGAGCCTT	896
repeat_region	/rpt family="L1MB3A"	40900.41359	Db	1948		AGGCAGCTTAGGCCCTGGGAGCCAGTCCACTGAGCAGCAAGGAGCTGAGTGAGCCTT	2007
repeat_region	/rpt family="L1MB3A"	41423.42036	Qy	897		TTGCAGGAAAGGCTAAGAAAGGAAACCAATCTTAAACACACACAAAGAACTGTC	956
repeat_region	/rpt family="L1MB3A"	42037.42253	Db	2008		TTGCAGGAAAGGCTAAGAAAGGAAACCAATCTTAAACACACACAAAGAACTGTC	2067
repeat_region	/rpt family="L1MB1"	complement(42651.43088)	Qy	957		ATGCTTTGGGAACTGTGTTTATTTGCTATATGGGTCCCCAAAATGGGTAACTAGACTT	1016
repeat_region	/rpt family="L2"	43091.43121	Db	2068		ATGCTTTGGGAACTGTGTTTATTTGCTATATGGGTCCCCAAAATGGGTAACTAGACTT	2127
repeat_region	/rpt family="polypurine"	94.7%; Score 1834.2; DB 9; Length 174445;	Qy	1017		CAGAGAGATGAGCAGAGCAAGGAGAAATCTGGCTGCTCTTCCATTTTCATCTCTGTT	1076
	Best Local Similarity 98.9%; Pred. No. 0;		Db	2128		CAGAGAGATGAGCAGAGCAAGGAGAAATCTGGCTGCTCTTCCATTTTCATCTCTGTT	2187
	Matches 1887; Conservative 0; Mismatches 13; Indels 8; Gaps 4;		Qy	1077		ATCTCAGGTGAGCTGGTAGAGGGAGACATTTAGAAAAAAATGAAAAACAAACAAATAC	1136
			Db	2188		ATCTCAGGTGAGCTGGTAGAGGGAGACATTTAGAAAAAAATGAAAAACAAACAAATAC	2247
			Qy	1137		TAATGAGTACCTGAGGCTGGGAGTCTTTGACTCCTCACTTAAATTCGTTTAGTGA	1196
			Db	2248		TAATGAGTACCTGAGGCTGGGAGTCTTTGACTCCTCACTTAAATTCGTTTAGTGA	2307
			Qy	1197		GAACCTTTCAATTTCTTTTATTTAGAGGCCAGCTTACTGTTGGTGGCAAAATGCCA	1256
			Db	2308		GAACCTTTCAATTTCTTTTATTTAGAGGCCAGCTTACTGTTGGTGGCAAAATGCCA	2367
			Qy	1257		ACATAAGTTAATAGAAAGTTGGCCAAATTTCAACCCCATTTTCTGTGGTTTGGGCTCCACAT	1316
			Db	2368		ACATAAGTTAATAGAAAGTTGGCCAAATTTCAACCCCATTTTCTGTGGTTTGGGCTCCACAT	2427
			Qy	1317		TGCAATGTTCAATGCCAGTCTGCTGACACCGAGGAGTACTAGCCAGCAGCAAAAGGC	1376
			Db	2428		TGCAATGTTCAATGCCAGTCTGCTGACACCGAGGAGTACTAGCCAGCAGCAAAAGGC	2487
			Qy	1377		AGGTAGCTGAATTTGCTTTCTGCTCTTTTACATTTCTTTTAAATAAGCATTTAGTGCTC	1436
			Db	2488		AGGTAGCTGAATTTGCTTTCTGCTCTTTTACATTTCTTTTAAATAAGCATTTAGTGCTC	2547
			Qy	1437		AGTCCCTACTGAGTACTCTTTCTCTCCCTCTCTGAAATTTAAATTTCAACTTGAAT	1496
			Db	2548		AGTCCCTACTGAGTACTCTTTCTCTCCCTCTCTGAAATTTAAATTTCAACTTGAAT	2607
			Qy	1497		TTGCAAGGATTACACATTTTCACTGTGATGATATTTGTTGTCAGNGGAAAAAGAAAGTG	1556
			Db	2608		TTGCAAGGATTACACATTTTCACTGTGATGATATTTGTTGTCAGNGGAAAAAGAAAGTG	2667
			Qy	1557		TCTTTGTTTAAATTTACTTGGTTTGTGAATCCATCTTGTCTTTTCCCATTTGGAACCTAGT	1616
			Db	2668		TCTTTGTTTAAATTTACTTGGTTTGTGAATCCATCTTGTCTTTTCCCATTTGGAACCTAGT	2727
			Qy	1617		CATTAAACCCATCTGAACTGGTAGAAAAACATCTGAGAGGAGTACTATCAGCATCTGA	1676

```

Db      2728  CATTACCCCTCTGACTGGTAGAAAAACATCTCGAAGAGCTAGTCTATCAGCATCTGA 2787
QY      1677  CAGGTGAATTTGGATGTTCTCAGAACCATTTCCACCAGACAGCCTGTTTCTATCTGCTTT 1736
Db      2788  CAGGTGAATTTGGATGTTCTCAGAACCATTTCCACCAGACAGCCTGTTTCTATCTGCTTT 2847
QY      1737  AATAAATAGTTTGGGTTTCTCATGTCATGATGATGATGATGATGATGATGATGATGATGAT 1796
Db      2848  AATAAATAGTTTGGGTTTCTCATGTCATGATGATGATGATGATGATGATGATGATGATGAT 2907
QY      1797  AGTCTGTGACTTCAAGTTTGTAGTCAGACCCCAACCAACCAACCAACCAACCAACCAACCA 1856
Db      2908  AGTCTGTGACTTCAAGTTTGTAGTCAGACCCCAACCAACCAACCAACCAACCAACCAACCA 2967
QY      1857  TGCAACATATGACTGTTTGTAAATTAAGTACCATGCTCTTTATATAA 1904
Db      2968  TGCAACATATGACTGTTTGTAAATTAAGTACCATGCTCTTTATATAA 3015

RESULT 11
LOCUS   AC022597
DEFINITION Homo sapiens chromosome 8 clone RP11-19A13 map 8, WORKING DRAFT
ACCESSION AC022597
VERSION   AC022597.7 GI:22123001
KEYWORDS  HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 184361)
          Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
          Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,P.,
          Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
          Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
          DeArellano,K., Dewar,K., Donino,M., Doyle,M., Fenestor,J.,
          Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
          Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
          Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
          Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
          MacDonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
          McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
          Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
          Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
          Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
          Stojanovic,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J.,
          Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
          Zimmer,A. and Zody,M.
          Direct Submission
          Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
          3 (bases 1 to 184361)
          Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
          Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
          Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
          Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
          Fero,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
          Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
          Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
          Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
          Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
          McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
          Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
          O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
          Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
          Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
          Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
          Testfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,

```

TITLE  
JOURNAL

## COMMENT

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.  
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 6, 2002 this sequence version replaced gi:14010820.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4195

Center clone name: 19\_A\_13

----- Summary Statistics

Sequencing vector: M13; M77815; 3% of reads

Sequencing vector: Plasmid; n/a; 97% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 180318 bases at least Q40

Consensus quality: 182555 bases at least Q30

Consensus quality: 183338 bases at least Q20

Insert size: 173000; agarose-fp

Insert size: 183761; sum-of-contigs

Quality coverage: 8.2 in Q20 bases; agarose-fp

Quality coverage: 7.8 in Q20 bas.

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1054: contig of 1054 bp in length  
\* 1055 1154: gap of 100 bp  
\* 1155 2825: contig of 1671 bp in length  
\* 2826 2925: gap of 100 bp  
\* 2926 29148: contig of 26223 bp in length  
\* 29149 59090: contig of 100 bp  
\* 59091 59190: gap of 100 bp  
\* 59191 110096: contig of 50906 bp in length  
\* 110097 110196: gap of 100 bp  
\* 110197 149304: contig of 39108 bp in length  
\* 149305 149404: gap of 100 bp  
\* 149405 184361: contig of 34957 bp in length.

## FEATURES

## source

1. 184361  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-19A13"  
/clone\_lib="RPC1-11 Human Male BAC"

## misc\_feature

1. 1054  
/note="assembly\_fragment"  
clone end:SP6  
vector side:left"

## misc\_feature

1155..2825  
/note="assembly\_fragment"

## misc\_feature

2926..29148  
/note="assembly\_fragment"

## misc\_feature

29249..59090  
/note="assembly\_fragment"

## misc\_feature

59191..110096  
/note="assembly\_fragment"

## misc\_feature

110197..149304  
/note="assembly\_fragment"

```
misc_feature 149405..184361
/note="assembly_fragment
clone end:T7
vector_side:right"
BASE COUNT 54282 a 40660 c 40193 g 48623 t 603 others
ORIGIN
Query Match 94.7%; Score 1832.6; DB 2; Length 184361;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1886; Conservative 0; Mismatches 14; Indels 8; Gaps 4;
QY 1 AATGGTATGCCAACTTAAGTATTTCAGGGTGGCCCAATAGAACAAAGATGCACTCGCTG 60
Db 64032 AATGGTATGCCAACTTAAGTATTTCAGGGTGGCCCAATAGAACAAAGATGCACTCGCTG 64091
QY 61 TGAATTTAAGACAAGCTGTATTAACAGAACTCCATGCAAGAGGNGGCGGGCAGGA 120
Db 64092 TGAATTTAAGACAAGCTGTATTAACAGAACTCCATGCAAGAGGNGGCGGGCAGGA 64151
QY 121 GAATCTCCGCTGTGTCAGACAGGGGCTTAAGAGGGTCTCCACACTGCTGTAGGGCT 180
Db 64152 GAATCTCCGCTGTGTCAGACAGGGGCTTAAGAGGGTCTCCACACTGCTGTAGGGCT 64211
QY 181 GTTGCAATTTTTATTTAGTAGAAGTGAAGGCGCTCTTCAACTTTTTTCCCTTGCGC 240
Db 64212 GTTGCAATTTTTATTTAGTAGAAGTGAAGGCGCTCTTCAACTTTTTTCCCTTGCGC 64271
QY 241 TGGAGAATTTAGAACTCAGAACTTCTCCCTCTCTTTTAAATTTTGTGCTCTTTTGCACAA 360
Db 64272 TGGAGAATTTAGAACTCAGAACTTCTCCCTCTCTTTTAAATTTTGTGCTCTTTTGCACAA 64331
QY 301 GAAGGCAACATAATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db 64332 GAAGGCAACATAATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 64391
QY 361 TTACTCACTAAAGGCTTCAATTTAGTCCAGATTTTGTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 64392 TTACTCACTAAAGGCTTCAATTTAGTCCAGATTTTGTGCTGCTGCTGCTGCTGCTGCTGCT 64451
QY 421 CCTGCTTATTTAGCCGAGATCTGGTC ---TTTTTNTGNTTTTTTTTTTTTTTTTTTTCGGTC 476
Db 64452 CCTGCTTATTTAGCCGAGATCTGGTC ---TTTTTNTGNTTTTTTTTTTTTTTTTTTTCGGTC 64511
QY 477 TCCCAAGAGCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 536
Db 64512 TCCCAAGAGCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 64570
QY 537 GCTAAAGACATGCAATTTTAAACTAGGCACTTCTTATTTCTTCTCTTTTAAATAACA 596
Db 64571 GCTAAAGACATGCAATTTTAAACTAG --CAACTCTTATTTCTTCTTTTAAATAACA 64628
QY 597 TAGCATTAATCCCAATCTTATTAAGACCTGACAGCTTGAGAAAGGTCACTACTGCAAT 656
Db 64629 TAGCATTAATCCCAATCTTATTAAGACCTGACAGCTTGAGAAAGGTCACTACTGCAAT 64688
QY 657 TTATAGGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 716
Db 64689 TTATAGGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 64748
QY 717 TTGATGACAGAGGTTAGAGGTATTTGATTTTCAAGAGGAAACACACGCGCAGAAAT 776
Db 64749 TTGATGACAGAGGTTAGAGGTATTTGATTTTCAAGAGGAAACACACGCGCAGAAAT 64808
QY 777 GAAGGGCCAGGCTTACTGAGGCTGCTCCAGTGGAGGCTCATGGTGGGACATGGAAAGA 836
Db 64809 GAAGGGCCAGGCTTACTGA --GCTGTCCAGTGGAGGCTCATGGTGGGACATGGAAAGA 64867
QY 837 AGGAGGCTTAGGCGCTGGGAGCCCACTGAGCAAGCAAGGCACTGAGTGAAGCCTT 896
Db 64868 AGGAGGCTTAGGCGCTGGGAGCCCACTGAGCAAGCAAGGCACTGAGTGAAGCCTT 64927
QY 897 TTGACGAAAAGGCTAAGAAAAGGAACCACTTCTTAAACACACAAAGAACTGTCCAA 956
Db 897 TTGACGAAAAGGCTAAGAAAAGGAACCACTTCTTAAACACACAAAGAACTGTCCAA 956
```

```
Db 64928 TTGACGAAAAGGCTAAGAAAAGGAAAACCAATTTCTTAAACACAAACAAGAACTGTCCAA 64987
QY 957 ATGCTTTTGGGAACTGTGTTTATTTGCTTATATATGCGTCCCAAAATGGGTAACTAGACTT 1016
Db 64988 ATGCTTTTGGGAACTGTGTTTATTTGCTTATATATGCGTCCCAAAATGGGTAACTAGACTT 65047
QY 1017 CACAGAGAATGACAGAGAGCAAGAGAAATCTGGCTGTCTCTTCCATTTTCACTTCTGTT 1076
Db 65048 CACAGAGAATGACAGAGAGCAAGAGAAATCTGGCTGTCTCTTCCATTTTCACTTCTGTT 65107
QY 1077 ATCTCAGGTGAGCTGTGTAGAGGGAGACATTTAGAAAAAATGAAAAAACAACAATTTAC 1136
Db 65108 ATCTCAGGTGAGCTGTGTAGAGGGAGACATTTAGAAAAAATGAAAAAACAACAATTTAC 65167
QY 1137 TAATGAGGTACGCTGAGGCTCTGGAGTCTCTTGTGACCTCACTACTTAATTCCTGTTAGTGA 1196
Db 65168 TAATGAGGTACGCTGAGGCTCTGGAGTCTCTTGTGACCTCACTACTTAATTCCTGTTAGTGA 65227
QY 1197 GAAACCTTTCAATTTTCTTTTATTAAGAGGGCCAGCTTACTGTGTCGCAAAATTGCCA 1256
Db 65228 GAAACCTTTCAATTTTCTTTTATTAAGAGGGCCAGCTTACTGTGTCGCAAAATTGCCA 65287
QY 1257 ACATAAGTTTAAATAGAAAGTTGCGCAATTTTCAACCCCAATTTCTGTGTTTGGGCTCCAAT 1316
Db 65288 ACATAAGTTTAAATAGAAAGTTGCGCAATTTTCAACCCCAATTTCTGTGTTTGGGCTCCAAT 65347
QY 1317 TGCATATGTTCAATGCCAGCTGCTGTGACACCGAGGAGTACTAGCCAGCACAAAAGGC 1376
Db 65348 TGCATATGTTCAATGCCAGCTGCTGTGACACCGAGGAGTACTAGCCAGCACAAAAGGC 65407
QY 1377 AGGGTAGCTCAATTTGCTTCTGCTCTTTTAAATTAAGCAATTTAGTGTCTC 1436
Db 65408 AGGGTAGCTCAATTTGCTTCTGCTCTTTTAAATTAAGCAATTTAGTGTCTC 65467
QY 1437 AGTCCCTACTGAGTACTCTTCTCTCCCTCTCTGAAATTTTAAATTTCAACTTTGCAACTTGAAT 1496
Db 65468 AGTCCCTACTGAGTACTCTTCTCTCCCTCTCTGAAATTTTAAATTTCAACTTTGCAACTTGAAT 65527
QY 1497 TTGCAAGATTTACACATTTCTGTGATATATTTGTTGTCAGNAGAAAAAGAAAGTG 1556
Db 65528 TTGCAAGATTTACACATTTCTGTGATATATTTGTTGTCAGNAGAAAAAGAAAGTG 65587
QY 1557 TCTTTGTTTAAAAATTTACTTGGTTTGTGAATCCATCTGCTGCTTTTCCCATTTGAACTAGT 1616
Db 65588 TCTTTGTTTAAAAATTTACTTGGTTTGTGAATCCATCTGCTGCTTTTCCCATTTGAACTAGT 65647
QY 1617 CATTAACCCATCTCTGAACTGGTAGAAAAACATCTGAAAGAGCTAGTCTATCAGCATCTGA 1676
Db 65648 CATTAACCCATCTCTGAACTGGTAGAAAAACATCTGAAAGAGCTAGTCTATCAGCATCTGA 65707
QY 1677 CAGGTGAATTTGGATGGTTCTCAGAACCATTTTACCAGACAGCTGTTTCTATCTGTTT 1736
Db 65708 CAGGTGAATTTGGATGGTTCTCAGAACCATTTTACCAGACAGCTGTTTCTATCTGTTT 65767
QY 1737 AATAAATTTAGTTTGGGTTCTCTACATGCAATAACAAACCTGCTCCCAATCTGTCACATAAA 1796
Db 65768 AATAAATTTAGTTTGGGTTCTCTACATGCAATAACAAACCTGCTCCCAATCTGTCACATAAA 65827
QY 1797 AGTCTGTGACTTGAAGTTTGTAGCAGACCCCAACCAAACTTTATTTTCTATGTGTTTTT 1856
Db 65828 AGTCTGTGACTTGAAGTTTGTAGCAGACCCCAACCAAACTTTATTTTCTATGTGTTTTT 65887
QY 1857 TGCAACATATGAGTGTGTTTGAATAAATAAGTACCAGTCTTCTTATTTAA 1904
Db 65888 TGCAACATATGAGTGTGTTTGAATAAATAAGTACCAGTCTTCTTATTTAA 65935
```

```
RESULT 12
AF247704
LOCUS
DEFINITION Homo sapiens homeobox protein NKX3.1 mRNA, complete cds.
ACCESSION AF247704
VERSION AF247704.1 GI:9963969
```



QY 1480 TTCTTTCAACTGCAATTCGCAAGGATTAACATTTCACTGTGATGATATATTTGTTGCA 1539  
Db |||||  
2865 TTCTTTCAACTGCAATTTGCAAGGTTTACATTTCACTGTGATGATATATTTGTTGCA 2924  
QY |||||  
1540 GNGAAGAAAAGAGTCTTTGTTTAAATTTACTTTGTTGTAATCCATCTTCTTTT 1599  
Db |||||  
2925 ---AAAAAAAAGAGTCTTTGTTTAAATTTACTTTGTTTGTGAATCCATCTTGC--TT 2979  
QY |||||  
1600 TCCCAATGGAATAGTCAATTAACCATCTCTGAACCTGTGTAAGAAAACATCTGAAGACT 1659  
Db |||||  
2980 TCCCAATGGAATAGTCAATTAACCATCTCTGAACCTGTGTAAGAAAACATCTGAAGACT 3039  
QY |||||  
1660 AGTCTATGAGACTGACAGAGTGAATGGAATGTTCTCAGAACCATTTTCAACCCAGACAGC 1719  
Db |||||  
3040 AGTCTATCAGCATCTGACAGTGAATGGAATGTTCTCAGAACCATTTTCAACCCAGACAGC 3099  
QY |||||  
1720 CTGTTTCTATCTCTGTTTAAATTTAGTCTGTTGTTCTCTACATGCATTAACAAACCTTGCT 1779  
Db |||||  
3100 CTGTTTCTATCTCTGTTTAAATTTAGTCTGTTGTTCTCTACATGCATTAACAAACCTTGCT 3159  
QY |||||  
1780 CCAATCTGTCACATAAAAGTCTGTGACTTGAAGTTTGTAGTCAGCACCCCCACCAACCTTTA 1839  
Db |||||  
3160 CCAATCTGTCACATAAAAGTCTGTGACTTGAAGTTTGTAGTCAGCACCCCCACCAACCTTTA 3219  
QY |||||  
1840 TTTTCTATGTTTCTGTTTGAACATATGAGTGTGTTTGAATAAAGT 1886  
Db |||||  
3220 TTTTCTATGTTTCTGTTTGAACATATGAGTGTGTTTGAATAAAGT 3266

RESULT 13  
AR244306/c  
LOCUS AR244306 2051 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 334 from patent US 6476207.  
ACCESSION AR244306  
VERSION AR244306.1 GI:27292093  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 2051)  
AUTHORS Zhang,J., Astle,J.H., Carroll,E. III, Endege,W.O., Ford,D.M.,  
Monahan,J.E., Schlegel,R. and Steinmann,K.E.  
TITLE Genes and gene expression products that are differentially  
regulated in prostate cancer  
JOURNAL Patent: US 6476207-A 334 05-NOV-2002;  
FEATURES Location/Qualifiers  
source  
1. 2051  
BASE COUNT 630 a 427 c 408 g 555 t 31 others  
ORIGIN

Query Match 71.7%; Score 1388.8; DB 6; Length 2051;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 1438; Conservative 1; Mismatches 12; Indels 5; Gaps 4;

QY 449 TTTTNTGNTNTTTTNTTTTCCGTCCTCCCAAGCTTTTATCTGTCCTTGACTTTTAA 508  
Db |||||  
1814 TGTTTTTTTTTTTTTTTTTTCGTCCTCCCAAGCTTTATCTGTCCTTGACTTTTAA 1755  
QY |||||  
509 AAAAGTTGGGGCAGATCTGAATGGCTAAAAGACATGCAATTTTAAACTAGGCAA 568  
Db |||||  
1754 AAAAGTTGGGGCAGATCTGAAT--GGCTAAAAGACATGCAATTTTAAWACTAG--CA 1698  
QY |||||  
569 CTCTTATTTCTTTTCCCTTTAAATATACATAGATTAATCCCAATCTCTATTTAAGACC 628  
Db |||||  
1697 ACTCTTATTTCTTTTAAATATACATAGATTAATCCCAATCTCTATTTAAGACC 1638  
QY |||||  
629 TGACGCTTGAGAGTCACTACTGCAATTAAGGACCTTCTGGTGGTTCTGCTTTACG 688  
Db |||||  
1637 TGACGCTTGAGAGGTCACACTACTGCAATTAAGGACCTTCTGGTGGTTCTGCTTTACG 1578  
QY |||||  
689 TTTGAAGTCTGCAATCTTTGAGATCTTTGATCAGAGGAGTGAAGGATTTGGAT 748  
Db |||||

Db 1577 TTTGAAGTCTGACAAATCCTTTGAAATCTTTTGCAATGACAGAGGATTAAGAGTATTGGATT 1518  
QY |||||  
749 TTCAACAGAGGAAAGAACACACAGCGCAGAAATGAAGGGCCAGGCTTACTGAGGCTGTCCAGTGG 808  
Db |||||  
1517 TTCAACAGAGGAAAGAACACACAGCGCAGAAATGAAGGGCCAGGCTTACTGA--GCTGTCCAGTGG 1459  
QY |||||  
809 AGGGCTCATGGGTGGGACATGGAAGAGAGGAGGAGGCTAGGCCCTGGGGAGCCAGTCCAC 868  
Db |||||  
1458 AGGGCTCATGGGTGGGACATGGAAGAGAGGAGGAGGCTAGGCCCTGGGGAGCCAGTCCAC 1399  
QY |||||  
869 TGAGCAAGCAAGGAGCTGAGTGAAGCTTTTGAGGAAAGGCTTAAGAAAGAGGAAAGCA 928  
Db |||||  
1398 TGAGCAAGCAAGGAGCTGAGTGAAGCTTTTGAGGAAAGGCTTAAGAAAGAGGAAAGCA 1339  
QY |||||  
929 TTTAAAAACAAACAGAAACTGTCCAAATGCTTTGGGAACTGTCTTTATTTCCTTATAT 988  
Db |||||  
1338 TTTAAAAACAAACAGAAACTGTCCAAATGCTTTGGGAACTGTCTTTATTTCCTTATAT 1279  
QY |||||  
989 GGGTCCCAAAATGGTAACTTACCTAGACTTCAGAGAAATGAGCAGAGCAAGAGGAAAT 1048  
Db |||||  
1278 GGGTCCCAAAATGGTAACTTACCTAGACTTCAGAGAAATGAGCAGAGCAAGAGGAAAT 1219  
QY |||||  
1049 CTGGCTGTCTTCCATTTTCAATTTCTGTTATCTCAGGTGAGTGGTAGAGGGGAGACATTA 1108  
Db |||||  
1218 CTGGCTGTCTTCCATTTTCAATTTCTGTTATCTCAGGTGAGTGGTAGAGGGGAGACATTA 1159  
QY |||||  
1109 GAAAAAATGAAACAAACAAACAAATTAATAGAGTACGCTGAGGCTGGGAGTCTCTT 1168  
Db |||||  
1158 GAAAAAATGAAACAAACAAACAAATTAATAGAGTACGCTGAGGCTGGGAGTCTCTT 1099  
QY |||||  
1169 GACTCCACTACTTAATTCGGTTTGTAGGAAACCTTTCAATTTCTTTTATTAGAAGGC 1228  
Db |||||  
1098 GACTCCACTACTTAATTCGGTTTGTAGGAAACCTTTCAATTTCTTTTATTAGAAGGC 1039  
QY |||||  
1229 CAGTTACTGTGTGGGCAAAATTTGCCAACATAAGTTAAATAGAAAGTTGGCAATTTCCAC 1288  
Db |||||  
1038 CAGTTACTGTGTGGGCAAAATTTGCCAACATAAGTTAAATAGAAAGTTGGCAATTTCCAC 979  
QY |||||  
1289 CCCATTTCTGTGTTGGGCTCCACATTTGCAATTTCAATGCCACGCTGTGCTGACACC 1348  
Db |||||  
978 CCCATTTCTGTGTTGGGCTCCACATTTGCAATTTCAATGCCACGCTGTGCTGACACC 919  
QY |||||  
1349 GACCGAGTACTAGCCAGCACAAAGGAGGCTAGCCTGAATTTCTTCTTCTCTTTTACA 1408  
Db |||||  
918 GACCGAGTACTAGCCAGCACAAAGGAGGCTAGCCTGAATTTCTTCTTCTCTTTTACA 859  
QY |||||  
1409 TTTCTTTTAAATAAGCATTTAGTGTCTGCTCCTACTGAGTACTCTTTCTCTCCCTCC 1468  
Db |||||  
858 TTTCTTTTAAATAAGCATTTAGTGTCTGCTCCTACTGAGTACTCTTTCTCTCCCTCC 799  
QY |||||  
1469 TCTGAATTTAATCTTTCAACTTGCATTTGCAAGGATTAACATTTCACTGTGATGTAT 1528  
Db |||||  
798 TCTGAATTTAATCTTTCAACTTGCATTTGCAAGGATTAACATTTCACTGTGATGTAT 739  
QY |||||  
1529 ATTGTGTTGAGNGAAAAGAAAAGTCTTTGTTTAAATTTACTTTGTTTCTGTAATCC 1588  
Db |||||  
738 ATTGTGTTGCA--AAAAAAAAGAGTCTTTGTTTAAATTTACTTTGTTTCTGTAATCC 680  
QY |||||  
1589 ATCTTGCTTTTCCCAATGGAACTAGTCAATTAACCATCTCTGAACCTGGTAGAAAACA 1648  
Db |||||  
679 ATCTTGCTTTTCCCAATGGAACTAGTCAATTAACCATCTCTGAACCTGGTAGAAAACA 620  
QY |||||  
1649 TCTGAAGAGTACTCTATCAGCATCTGACAGGTGAATTTGGATGTTCTCAGAACCAATTTC 1708  
Db |||||  
619 TCTGAAGAGTACTCTATCAGCATCTGACAGGTGAATTTGGATGTTCTCAGAACCAATTTC 560  
QY |||||  
1709 ACCCAGACGCTGTTTCTATCTGTTTAAATTTAGTGTGTTTCTTACATGATATAA 1768  
Db |||||  
559 ACCCAGACGCTGTTTCTATCTGTTTAAATTTAGTGTGTTTCTTACATGATATAA 500  
QY |||||  
1769 CAAACCTGCTCCAAATCTGTACATAAAAGTCTGTGACTTTGAAGTTTCTGACGACCCCC 1828  
Db |||||  
499 CAAACCTGCTCCAAATCTGTACATAAAAGTCTGTGACTTTGAAGTTTCTGACGACCCCC 440

Qy	1829	ACAAACTTTATTTTTCTAIGTGTGTTTTTGCAACATAAGTAGTGTGGTAAAAATAAGTAC	1889
Dd	439	ACCAAATTTATTTTTCTATGTGTTTTTGCAACATATGAGTGTTCGAAAATAAGTAC	380
Qy	1889	CCATGTCCTTTATTAAA	1904
Dd	379	CCAATGCTTTATTAGA	364
RESULT 14			
AR244053/c			
LOCUS AR244053 1024 bp DNA linear PAT 20-DEC-2002			
DEFINITION Sequence 81 from patent US 6476207.			
ACCESSION AR244053			
VERSION AR244053.1 GI:27291840			
KEYWORDS			
SOURCE Unknown.			
ORGANISM Unclassified.			
REFERENCE 1 (bases 1 to 1024)			
AUTHORS Zhang,J., Astle,J.H., Carroll,E. III, Endege,W.O., Ford,D.M., Monahan,J.E., Schlegel,R. and Steinmann,K.E.			
TITLE Genes and gene expression products that are differentially regulated in prostate cancer			
JOURNAL Patent: US 6476207-A 81 05-NOV-2002;			
FEATURES Location/Qualifiers			
source 1..1024			
BASE COUNT 248 a 239 c 210 g 281 t 46 others			
ORIGIN			
Query Match 33.1%; Score 641.2; DB 6; Length 1024;			
Best Local Similarity 97.3%; Pred. No. 5.2e-145;			
Matches 681; Conservative 0; Mismatches 16; Indels 3; Gaps 3;			
Qy	448	CTTTTTNTGNTTTTTTTTTTTTCCGTCTCCCCAAGCTTTATCTGCTTGACTTTTA	507
Dd	799	CTCTTNNNNNTTTTTTTTTTCNGTCTCCCCAAGCTTTATCTGCTTGACTTTNA	740
Qy	508	AAAAAGTTTGGGGCAGATCTGAATTTGGGCTAAAAACATGCATTTTTAAAACTAGCCA	567
Dd	739	AAAAAGTTTGGGGCAGATCTGAATTTGGTAAAGACATGCATTTTTAAAACTA-GCA	682
Qy	568	ACTTCTATTCTTCTTTTAAAAATACATAGCATTAATCCCAGAACCTTATTTAAAGAC	627
Dd	681	ACTTCTATTCTTCTTTTAAAAATACATAGCATTAATCCCAGAACCTTATTTAAAGAC	622
Qy	628	CTCACAGCTTGAGAGGTCACTACTGCATTTATAGGACCCTTCTGGTGGTCTCGCTGTTC	687
Dd	621	CTCACAGCTTGAGAGGTCACTACTGCATTTATAGGACCCTTCTGGTGGTCTCGCTGTTC	562
Qy	688	GTTTGAAGCTTGACAATCTCTTGAGAACTCTTTGCAATGCAGAGGAGTTAAGAGGTATTCGAT	747
Dd	561	GTTTGAAGCTTGACAATCTCTTGAGAACTCTTTGCAATGCAGAGGAGTTAAGAGGTATTCGAT	502
Qy	748	TTTTCACAGAGGAAGAACACACGCCAGAACTGAAGGGCCAGGCTTACTCAGAGCTGTCCAGTG	807
Dd	501	TTTTCACAGAGGAAGAACACACGCCAGAACTGAAGGGCCAGGCTTACTCAGAGCTGTCCAGTG	443
Qy	808	GAGGGCTCATGGGTGGGACATGGAAGAGAGGAGCCCTAGGCCCTGGGGAGCCAGTCCA	867
Dd	442	GAGGGCTCATGGGTGGGACATGGAAGAGAGGAGCCCTAGGCCCTGGGGAGCCAGTCCA	383
Qy	868	CTGAGCAGCAAGGACCTGATGAGCCCTTTTGAGGAAAAAGGCTTAAGAAAAAGGAAAAACC	927
Dd	382	CTGAGCAAGCAAGGACCTGATGAGCCCTTTTGAGGAAAAAGGCTTAAGAAAAAGGAAAAACC	323
Qy	928	ATTTCTAAAAACACAAAGAACTGTCGAAATGCTTTTGGGAACCTGTTTATTCGCTATAA	987
Dd	322	ATTTCTAAAAACACAAAGAACTGTCGAAATGCTTTTGGGAACCTGTTTATTCGCTATAA	263
Qy	988	TGGGTCCCCAAAATGGGTAACTAGACTTCAGAGAGAAATGAGCAGAGCAAGGAGGAGAA	1047

Db	262	TGGGTCCCCAAAATGGGTAAACCTAGACTTTCAGAGAGAAATGAGCAGAGCAAAAGGAGAAA	203
Qy	1048	TCTGGCTGTCTCTCCATTTTTCATTTCTGTGTATCTCAGCTGAGCTGGTAGAGGGGAGACATT	1107
Db	202	TCTGGCTGTCTCTCCATTTTTCATTTCTGTGTATCTCAGCTGAGCTGGTAGAGGGGAGACATT	143
Qy	1108	AGAAAAAATGAACAACAACAACAATACTTAATGAGGTAC	1147
Db	142	AGAAAAAATGAACAACAACAACAATACTTAATGAGGTAC	103
RESULT 15			
LOCUS	AR244068	1024 bp	DNA linear PAT 20-DEC-2002
DEFINITION	Sequence 96 from patent US 6476207.		
ACCESSION	AR244068		
VERSION	AR244068.1	GI-27291855	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1024)		
AUTHORS	Zhang,J., Astle,J.H., Carroll,E. III, Endege,W.O., Ford,D.M., Monahan,J.E., Schlegel,R. and Steinmann,K.E.		
TITLE	Genes and gene expression products that are differentially regulated in prostate cancer		
JOURNAL	Patent: US 6476207-A 96 05-NOV-2002;		
FEATURES	Location/Qualifiers		
source	1..1024		
BASE COUNT	249 a	250 c	214 g
ORIGIN	249 a	250 c	214 g
Query Match	32.5%;	Score 628.4;	DB 6; Length 1024;
Best Local Similarity	37.4%;	Pred. No. 6.7e-142;	
Matches	678; Conservative	0; Mismatches	13; Indels 5; Gaps 4;
Qy	453	TTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTCGCTCCCCAAAGCTTTATCTGT	-CTTGACTTTTTTAAAAA 511
Db	795	TTTTTNNNNTTTTTTTTTTCNGCTCCCCAAAGCTTTATCGGTCCTTGACTTTTNAAAA	736
Qy	512	AGTTTGGGGCAGATTCTGAAATTGGGCTAAAAGACATGCATTTTAAAACTAGGCAACTT	571
Db	735	AGTTTGGGGCAGATTCTGAAATT-GGTTAAAAGACATGCATTTTAAAACTAG--CAATT	679
Qy	572	CTTATTTCTTTCTTTAAAATAATACATGCAATTAATCCCAATCCTATTAAAGACCTGA	631
Db	678	CTTATTTCTTTCTTTAAAATAATACATGCAATTAATCCCAATCCTATTAAAGACCTGA	619
Qy	632	CAGCTTGAGAAAGTCACTACTGCATTTATAGGACCTTCTGGTGGTTCTGCTGTACGTTT	691
Db	618	CAGCTTGAGAAAGTCACTACTGCATTTATAGGACCTTCTGGTGGTTCTGCTGTACGTTT	559
Qy	692	GAAGTCTGACAAATCCTTGTGAGAACTTTTGCATGCAGAGGAGGTAAGAGGTAATTCGATTTTC	751
Db	558	GAAGTCTGACAAATCCTTGTGAGAACTTTTGCATGCAGAGGAGGTAAGAGGTAATTCGATTTTC	499
Qy	752	ACAGAGGAACAACACAGCGCAGAAATGAAGGGCCAGGCTTACTGAGGCTGTCCAGTGGAGG	811
Db	498	ACAGAGGAAGAACAACAGCGCAGAAATGAAGGGCCAGGCTTACTGA-GCTGTCAGTGGAGG	440
Qy	812	GCTCATGGGTGGACATGGAAGAGAGGACGCTTAGCCCTCGGGAGCCAGTCCACTGA	871
Db	439	GCTCATGGGTGGGACATGGAAGAGAGGACGCTTAGCCCTCGGGAGCCAGTCCACTGA	380
Qy	872	GCAAGCAAGGCACTGAGTGAAGCCCTTTTGCAGGAAAAAGGCTTAAGAAAAAGGAAAAACCATTC	931
Db	379	GCAAGCAAGGCACTGAGTGAAGCCCTTTTGCAGGAAAAAGGCTTAAGAAAAAGGAAAAACCATTC	320
Qy	932	TAAAAACAACAAGAACTGCTCCAAATGCTTTGGGAACTGTGTTTATTCGCTTAATATGGG	991
Db	319	TAAAAACAACAAGAACTGCTCCAAATGCTTTGGGAACTGTGTTTATTCGCTTAATATGGG	260

```
Qy 992 TCCCCAAAATGGTAACCTAGACTTCAGAGAGATGACAGAGCAAGGAGAAATCTG 1051
Db 259 TCCCCAAAATGGTAACCTAGACTTCAGAGAGATGACAGAGCAAGGAGAAATCTG 200
Qy 1052 GCTGTCCTTCCATTTTCATTCTGTTATCTCAGGTGAGCTGTAGAGGGGAGACATTAGAA 1111
Db 199 GCTGTCCTTCCATTTTCATTCTGTTATCTCAGGTGAGCTGTAGAGGGGAGACATTAGAA 140
Qy 1112 AAAAATGAAACAACAACAATTAATGAGGTAC 1147
Db 139 AAAAATGAAACAACAACAATTAATGAGGTAC 104
```

Search completed: August 26, 2003, 17:06:58  
Job time : 9204 secs

	1	1930	99.7	1936	21	AAZ45667	Nucleotide sequenc
1	1836.4	94.9	2984	21	AAA06564	Human immunogeni	
2	1836.4	94.9	2984	22	AAAG3773	Human prostate CDN	
3	1836.4	94.9	2984	21	AAH93680	Human prostate-spe	
4	1836.4	94.9	2984	22	AAH84894	Human prostate-spe	
5	1836.4	94.9	2984	22	AAH02745	Prostate tumour an	
6	1836.4	94.9	2984	22	AAF86953	Human F705p invent	
7	1836.4	94.9	2984	22	AAFL95144	Human F705p cDNA s	
8	1836.4	94.9	2984	24	ABL95144		

Claim 6; Page 22-23; 30pp; English.

PS The present sequence represents a human prostate specific gene (PSG)  
XX (gene ID 244673). The mRNA levels for PSG are about 20 fold higher  
CC in the prostate than in other tissues. The PSG mRNA was found to be  
CC overexpressed in prostate tumour samples. The specification describes  
CC a method for diagnosing the presence of, or metastatic potential of,  
CC prostate cancer in a patient. The method comprises measuring PSG levels  
CC in a cell, tissue or bodily fluid sample of the patient and a control  
CC (i.e. a normal human without cancer), where increased PSG levels in  
CC the patient compared to the control is associated with the presence of,  
CC or metastasis of prostate cancer. The method can also be used for  
CC staging prostate cancer in a patient. The methods are used to detect,  
CC monitor, stage and give a prognosis for prostate cancer.  
XX  
SQ Sequence 1936 BP; 546 A; 392 C; 404 G; 588 T; 6 other;

Query Match 99.7%; Score 1930; DB 21; Length 1936;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTATGCCAACTTAAGTATTTACAGGGTGGCCCAATAGACAGATGCCTGCTG 60  
DB 1 AATGTATGCCAACTTAAGTATTTACAGGGTGGCCCAATAGACAGATGCCTGCTG 60  
QY 61 TGAATTTAAGCAAGCTGTATAAACAAGAACTCCACTGCAAGAGGGGCGGCCAGGA 120  
DB 61 TGAATTTAAGCAAGCTGTATAAACAAGAACTCCACTGCAAGAGGGGCGGCCAGGA 120  
QY 121 GAATTCGGCTGTCCAGACAGGGCTTAAGAGGGTCTCCACATGTCTGCTAGGGGCT 180  
DB 121 GAATTCGGCTGTCCAGACAGGGGCTTAAGAGGGTCTCCACATGTCTGCTAGGGGCT 180  
QY 181 GTTGCATTTTATTATAGTAGAAGTGAAGAGGCTCTCTCAACTTTTTCCTTGGGC 240  
DB 181 GTTGCATTTTATTATAGTAGAAGTGAAGAGGCTCTCTCAACTTTTTCCTTGGGC 240  
QY 241 TGGAGATTTAGATACAGAGTTTCTGGAGTTTTCAGGCTATCATATATACATCTCT 300  
DB 241 TGGAGATTTAGATACAGAGTTTCTGGAGTTTTCAGGCTATCATATATACATCTCT 300  
QY 301 GAAAGGCAACATAATCTCTCCCTCCCTTTTAAATTTTGTGCTTTTGGAGAA 360  
DB 301 GAAAGGCAACATAATCTCTCCCTCCCTTTTAAATTTTGTGCTTTTGGAGAA 360  
QY 361 TTACTCACTAAGGGCTCAATTTAGTCCAGATTTTGTGCTGGCTGCACCTAATATG 420  
DB 361 TTACTCACTAAGGGCTCAATTTAGTCCAGATTTTGTGCTGGCTGCACCTAATATG 420  
QY 421 CCTCGCTATTAGCCGAGATCTGCTCTTTTNTGNTTTTNTTNTTNTTNTTNTTNT 480  
DB 421 CCTCGCTATTAGCCGAGATCTGCTCTTTTNTGNTTTTNTTNTTNTTNTTNTTNT 480  
QY 481 CAAAGCTTTATCTGTCTGACTTTTAAAGTTTGGGGGAGATCTGAAATTTGGGCTA 540  
DB 481 CAAAGCTTTATCTGTCTGACTTTTAAAGTTTGGGGGAGATCTGAAATTTGGGCTA 540  
QY 541 AAAGACATGCAATTTTAAACTAGGCAACTCTTATTTCTTCTTTAAAAATACATAGC 600  
DB 541 AAAGACATGCAATTTTAAACTAGGCAACTCTTATTTCTTCTTTAAAAATACATAGC 600  
QY 601 ATTAAATCCCAATCTTATTAAAGACCTGACAGCTTGAGAGGTCACCTACGCAATTTAT 660  
DB 601 ATTAAATCCCAATCTTATTAAAGACCTGACAGCTTGAGAGGTCACCTACGCAATTTAT 660  
QY 661 AGACCTTCTGGGTCTGCTGTTACGTTTGAAGTCTGACAATCTCTGAGAACTTTTTC 720  
DB 661 AGACCTTCTGGGTCTGCTGTTACGTTTGAAGTCTGACAATCTCTGAGAACTTTTTC 720  
QY 721 ATCAGAGAGGTTAAGAGGTATGGATTTTTCAGAGGAAGAACACAGCGCAGAAATGAAG 780  
DB 721 ATCAGAGAGGTTAAGAGGTATGGATTTTTCAGAGGAAGAACACAGCGCAGAAATGAAG 780

QY 781 GGCCAGGCTTACTAGGCTGTCCAGTGGAGGGCTCATGGGTGGACATCGAAAAAGAGGC 840  
DB 781 GGCCAGGCTTACTAGGCTGTCCAGTGGAGGGCTCATGGGTGGACATCGAAAAAGAGGC 840  
QY 841 AGCTAGGCTTGGGGAGCCAGTCCACAGCAGCAGGAGGAGTGAAGCCTTTTTC 900  
DB 841 AGCTAGGCTTGGGGAGCCAGTCCACAGCAGCAGGAGGAGTGAAGCCTTTTTC 900  
QY 901 AGGAAAGGCTTAAGAAAAAGGAAACCAATTTAAACACAAAGAACTGTCCAAATGC 960  
DB 901 AGGAAAGGCTTAAGAAAAAGGAAACCAATTTAAACACAAAGAACTGTCCAAATGC 960  
QY 961 TTTGGGAAGCTGTATTATGGTCCCTATATGGTCCCAAAATGGGTAACTAGCTTCAG 1020  
DB 961 TTTGGGAAGCTGTATTATGGTCCCTATATGGTCCCAAAATGGGTAACTAGCTTCAG 1020  
QY 1021 GAGATGAGCAGAGCAAGAGGAAATCTGGCTGTCTCTTCCATTTTCACTCTGTATCT 1080  
DB 1021 GAGATGAGCAGAGCAAGAGGAAATCTGGCTGTCTCTTCCATTTTCACTCTGTATCT 1080  
QY 1081 CAGGTGAGCTGTAGAGGGGAGACATTAGAAAAAATGAAACAAACAACTACTAAT 1140  
DB 1081 CAGGTGAGCTGTAGAGGGGAGACATTAGAAAAAATGAAACAAACAACTACTAAT 1140  
QY 1141 GAGGTACGCTGAGGCTCGGAGTCTCTTACTCCACTACTTAATTTCCGTTAGTGAGAA 1200  
DB 1141 GAGGTACGCTGAGGCTCGGAGTCTCTTACTCCACTACTTAATTTCCGTTAGTGAGAA 1200  
QY 1201 CTTTCAATTTTCTTTATTAGAGGGCCAGCTTACTGTGTGGCAAAATTTGCCAAT 1260  
DB 1201 CTTTCAATTTTCTTTATTAGAGGGCCAGCTTACTGTGTGGCAAAATTTGCCAAT 1260  
QY 1261 AAGTTAATAGAAAGTTGGCCAAATTTACCCCAATTTTCTGTGGTGGCTCCACATTGCA 1320  
DB 1261 AAGTTAATAGAAAGTTGGCCAAATTTACCCCAATTTTCTGTGGTGGCTCCACATTGCA 1320  
QY 1321 ATGTTCAATGCCAGCTGTCTGCTGACACCGAGCGAGTACTAGCCAGACAAAGCGAGG 1380  
DB 1321 ATGTTCAATGCCAGCTGTCTGCTGACACCGAGCGAGTACTAGCCAGACAAAGCGAGG 1380  
QY 1381 TAGCCTGAATTTGCTTCTGCTCTTTAATTTTAAATTAAGCAATTTAGTGTCTAGTC 1440  
DB 1381 TAGCCTGAATTTGCTTCTGCTCTTTAATTTTAAATTAAGCAATTTAGTGTCTAGTC 1440  
QY 1441 CCTACTGAGTACTCTTCTCTCCCTCTCTGAAATTTAATTTTCTTCACTTGAATTTGC 1500  
DB 1441 CCTACTGAGTACTCTTCTCTCCCTCTCTGAAATTTAATTTTCTTCACTTGAATTTGC 1500  
QY 1501 AAGATTACACATTTCACTGTGATGTATATTGTGTCAGNGAAAGAAAAAGTGTCTT 1560  
DB 1501 AAGATTACACATTTCACTGTGATGTATATTGTGTCAGNGAAAGAAAAAGTGTCTT 1560  
QY 1561 TGTTTAAATTTACTGTTTGTGAATCTCTTGTCTTTTCCCATTTGGAACTAGTCAAT 1620  
DB 1561 TGTTTAAATTTACTGTTTGTGAATCTCTTGTCTTTTCCCATTTGGAACTAGTCAAT 1620  
QY 1621 AACCCATCTCTGAATCTGTAGAAAAACATCTGAAGAGCTAGTCTATCAGCATCTGACAGG 1680  
DB 1621 AACCCATCTCTGAATCTGTAGAAAAACATCTGAAGAGCTAGTCTATCAGCATCTGACAGG 1680  
QY 1681 TGAATTTGGATGTTCTCAGAACCAATTTCAAGAGAGCTGTTCTATCTCTGTTTAAATA 1740  
DB 1681 TGAATTTGGATGTTCTCAGAACCAATTTCAAGAGAGCTGTTCTATCTCTGTTTAAATA 1740  
QY 1741 AATTTAGTTTGGTCTCTACATGCAATAACAAACCTCTCTCAATCTGTCAATAAAGTC 1800  
DB 1741 AATTTAGTTTGGTCTCTACATGCAATAACAAACCTCTCTCAATCTGTCAATAAAGTC 1800  
QY 1801 TGTGACTTGAAGTTTGTAGTACAGCCCCCACTTTTATTTTCTATGTTTGTGCA 1860  
DB 1801 TGTGACTTGAAGTTTGTAGTACAGCCCCCACTTTTATTTTCTATGTTTGTGCA 1860  
QY 1861 ACATATGAGTGTGTTGAAAAATAAGTACCCTCTTTTAAAAAANAAGGGGC 1920

```

Db      1861 ACATATGAGTGTGTTGAAATAAGTACCCATGCTTTATTAATAAATAAATAAAGGCG 1920
QY      1921 GGCCGCGGACTAGTGA 1936
Db      1921 GGCCGCGGACTAGTGA 1936

RESULT 2
AAA06564
ID AAA06564 standard; cDNA; 2984 BP.
AC AAA06564;
XX
XX
DT 13-JUN-2000 (first entry)
XX
DE Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:335.
XX
XX Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW immunogenic; cytostatic; vaccine; ss.
XX
XX Homo sapiens.
XX
PN WO200004149-A2.
XX
PD 27-JAN-2000.
XX
XX 14-JUL-1999; 99WO-US15838.
XX
XX 14-JUL-1998; 98US-0115453.
XX 14-JUL-1998; 98US-0116134.
XX 23-SEP-1998; 98US-0159812.
XX 23-SEP-1998; 98US-0159822.
XX 15-JAN-1999; 99US-0232149.
XX 15-JAN-1999; 99US-0232880.
XX 09-APR-1999; 99US-0288946.
XX
XX (CORI-) CORIXA CORP.
XX
XX Dillon DC, Harlocker SL, Yuqiu J, Xu J, Mitcham JL;
XX WPI; 2000-171268/15.
XX
XX
XX New polypeptide useful for treating and diagnosing prostate cancer
XX comprises an immunogenic portion of prostate tumor protein -
XX
XX Claim 50; Page 207-208; 263pp; English.
XX
XX The present invention describes isolated polypeptides, comprising an
XX immunogenic portion of a prostate tumour protein (pnp). The polypeptides
XX and polynucleotides encoding them have cytostatic activity and can be
XX used in vaccines and in gene therapy. The polypeptides and
XX polynucleotides encoding them, antigen presenting cells which express
XX the polypeptides, antibodies against the polypeptides and vaccines
XX comprising them can be used for inhibiting the development of prostate
XX cancer in a patient. The polypeptides can be used to generate antibodies
XX or anti-idiotypic antibodies for passive immuno therapy. A portion of
XX the polynucleotides encoding the polypeptides can be used as a probe or
XX to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
XX AAA82000 to AAA82020 represent sequences used in the exemplification of
XX the present invention.
XX
XX Sequence 2984 BP; 837 A; 659 C; 645 G; 842 T; 1 other;
XX

Query Match          94.9%; Score 1836.4; DB 21; Length 2984;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1897; Conservative 0; Mismatches 12; Indels 7; Gaps 5;

QY      1 AATGGTATGCCAACTTAAGTATTTACAGGTGGCCCAATAGAACAAAGATGCACCTGCTG 60
Db      1067 AATGGTATGCCAACTTAAGTATTTACAGGTGGCCCAATAGAACAAAGATGCACCTGCTG 1126
QY      61 TGATTTTAAGACAAGCTGTATATAACAGAACTCCACTGCAAGGGGCGGCCGAGGA 120

```

```

Db      1127 TGATTTTAAGACAAGCTGTATATAACAGAACTCCACTGCAAGAGGGGGCCGCGCCAGGA 1186
QY      121 GAATCTCCGCTTGTCCAAGACAGGGGCCCTAAAGAGGGGTCTCCACACTGCTGCTAGGGGCT 180
Db      1187 GAATCTCCGCTTGTCCAAGACAGGGGCCCTAAAGAGGGGTCTCCACACTGCTGCTAGGGGCT 1246
QY      181 GTTGCAATTTTTTATTAGTAGAAGTGGAAAGCCCTCTTCTCAACTTTTTTCCCTTTGGGC 240
Db      1247 GTTGCAATTTTTTATTAGTAGAAGTGGAAAGCCCTCTTCTCAACTTTTTTCCCTTTGGGC 1306
QY      241 TCGAGAAATTTAGAAATCAGAAAGTTTCTCGAGTTTTCAGGCTATCATATATACCTATCTCT 300
Db      1307 TCGAGAAATTTAGAAATCAGAAAGTTTCTCGAGTTTTCAGGCTATCATATATACCTATCTCT 1366
QY      301 GAAAGGCAACATAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db      1367 GAAAGGCAACATAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1426
QY      361 TTACTCAGCTAAAGGGCTTCATTTTAGTCCAGATTTTTTAGTCTGGCTGCACCTAACTTATG 420
Db      1427 TTACTCAGCTAAAGGGCTTCATTTTAGTCCAGATTTTTTAGTCTGGCTGCACCTAACTTATG 1486
QY      421 CCTCGCTTATTAGCCCGAGATCTGGTCTTTTTTGTGTTTTTTTTTTTTTTTTTTTTTCCGCTCCC 480
Db      1487 CCTCGCTTATTAGCCCGAGATCTGGTCTTTTTT - TTTTTTTTTTTTTTTTTTCCGCTCCC 1544
QY      481 CAAAGCTTTATCTGCTTGACTTTTTTAAAAAGTTTGGGGCGAGATTCGAAATTTGGGCTA 540
Db      1545 CAAAGCTTTATCTGCTTGACTTTTTTAAAAAGTTTGGGGCGAGATTCGAAATTTGGGCTA 1603
QY      541 AAAGCATGCACTTTTAAAACTAGGCACTTCTTATTTCTTTCTTTAAAAATACATAGC 600
Db      1604 AAAGCATGCACTTTTAAAACTAG - CAACTCTTATTTCTTTCTTTAAAAATACATAGC 1661
QY      601 ATTAATCCAAATCCTATTAAAGACCTGACAGCTTGAGAAAGTGCACCTACTGCACTTTAT 660
Db      1662 ATTAATCCAAATCCTATTAAAGACCTGACAGCTTGAGAAAGTGCACCTACTGCACTTTAT 1721
QY      661 AGGACCTTCTGGTGTCTGCTGTTTACAGTCTGACAACTCTTCAGAAATCTTTTGC 720
Db      1722 AGGACCTTCTGGTGTCTGCTGTTTACAGTCTGACAACTCTTCAGAAATCTTTTGC 1781
QY      721 ATGACAGAGGTAAGAGGTATTGGATTTTCAAGAGGAAAGAACACAGCGCAGAAATGAAG 780
Db      1782 ATGACAGAGGTAAGAGGTATTGGATTTTCAAGAGGAAAGAACACAGCGCAGAAATGAAG 1841
QY      781 GGCAGGCTTACTGAGGCTGTCCAGTGGAGGCTCATGGTGGGACATGGAAAAAGAGGC 840
Db      1842 GGCAGGCTTACTGA - GCTGTCCAGTGGAGGCTCATGGTGGGACATGGAAAAAGAGGC 1900
QY      841 AGCCTAGGCCCTGGGAGCCCACTGAGCAAGCAAGGAGCTGAGTGAGCCCTTTTGC 900
Db      1901 AGCCTAGGCCCTGGGAGCCCACTGAGCAAGCAAGGAGCTGAGTGAGCCCTTTTGC 1960
QY      901 AGCAAAAGCTTAAGAAAAAGGAAACCAATTTTAAACACAACAAGAAACTGTCCAAATGC 960
Db      1961 AGCAAAAGCTTAAGAAAAAGGAAACCAATTTTAAACACAACAAGAAACTGTCCAAATGC 2020
QY      961 TTTGGAACTGTGTTTATGTCCTATAATGGTGTCCCAAAATGGGTAACTAGACTTCAGA 1020
Db      2021 TTTGGAACTGTGTTTATGTCCTATAATGGTGTCCCAAAATGGGTAACTAGACTTCAGA 2080
QY      1021 GAGATGACGAGAGCAAGGAGAAATCTGGCTGCTCTTCCATTTTCTTCTTCTTCTTCT 1080
Db      2081 GAGATGACGAGAGCAAGGAGAAATCTGGCTGCTCTTCCATTTTCTTCTTCTTCTTCT 2140
QY      1081 CAGGTGAGCTGGTAGGGGAGACATTAGAAAAAATAAACAACAACAACAATTAAT 1140
Db      2141 CAGGTGAGCTGGTAGGGGAGACATTAGAAAAAATAAACAACAACAACAATTAAT 2200
QY      1141 GAGGTGAGCTGGAGGCTGGAGTCTCTTGACTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200

```

Db 2201 GAGGTACGCTGAGGCGCTGGAGTCTCTTCACTCCACTTAATTCGGTTAGTCAGAAA 2260  
 QY 1201 CTTTCAATTTCTTTTATAGAGGCGCAGCTTACTGTGTGTCGCAAAAATGTCACAAAT 1260  
 Db 2261 CTTTCAATTTCTTTTATAGAGGCGCAGCTTACTGTGTGTCGCAAAAATGTCACAAAT 2320  
 QY 1261 AAGTTAATAAGATTTGGCCAAATTTCAACCCCAATTTCTGTGTTGGGCTCCACATTCGA 1320  
 Db 2321 AAGTTAATAAGATTTGGCCAAATTTCAACCCCAATTTCTGTGTTGGGCTCCACATTCGA 2380  
 QY 1321 ATGTTCAATGCCAGTCTGTCTGACACCGAGTACTAGCCAGCACAAGAGCGAGG 1380  
 Db 2381 ATGTTCAATGCCAGTCTGTCTGACACCGAGTACTAGCCAGCACAAGAGCGAGG 2440  
 QY 1381 TAGCCCTGAATTCCTTCTCTCTTTTACATTTCTTTTAAATAAGCATTTAGTCTCAGTC 1440  
 Db 2441 TAGCCCTGAATTCCTTCTCTCTTTTACATTTCTTTTAAATAAGCATTTAGTCTCAGTC 2500  
 QY 1441 CCTACTGAGTACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500  
 Db 2501 CCTACTGAGTACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2560  
 QY 1501 AAGGATTACACATTTCACTGTGATGATATATTTGCTGTCAGNGAAGAAAAGTGTCTT 1560  
 Db 2561 AAGGATTACACATTTCACTGTGATGATATATTTGCTGTCAGNGAAGAAAAGTGTCTT 2619  
 QY 1561 TGTTTAAATTAATCTGTTGTTGTAATCCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1620  
 Db 2620 TGTTTAAATTAATCTGTTGTTGTAATCCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 2679  
 QY 1621 AACCACCTCTGAACTGTAGAAAACATCTGAAGAGCTAGTCTATCAGCATCTGCACAG 1680  
 Db 2680 AACCACCTCTGAACTGTAGAAAACATCTGAAGAGCTAGTCTATCAGCATCTGCACAG 2739  
 QY 1681 TGAATTTGATGGTTCTCAGAACCAATTTCAACCCAGACAGCTGTTCTATCTCTGTTTAA 1740  
 Db 2740 TGAATTTGATGGTTCTCAGAACCAATTTCAACCCAGACAGCTGTTCTATCTCTGTTTAA 2799  
 QY 1741 AATTAGTTTGGTTCTCTACATGATATAACAAACCTGCTCCAATCTGTACATAAAGTC 1800  
 Db 2800 AATTAGTTTGGTTCTCTACATGATATAACAAACCTGCTCCAATCTGTACATAAAGTC 2859  
 QY 1801 TGTGACTTGAAGTTTGTAGTCAGACCCCAACCACTTATTTTCTATCTGTTTGTGCA 1860  
 Db 2860 TGTGACTTGAAGTTTGTAGTCAGACCCCAACCACTTATTTTCTATCTGTTTGTGCA 2919  
 QY 1861 ACATATGAGTGTGTTGAAAATAAAGTACCCATGCTCTTTTATTAATAAANAANAANA 1916  
 Db 2920 ACATATGAGTGTGTTGAAAATAAAGTACCCATGCTCTTTTATTAATAAANAANAANA 2975

## RESULT 3

AAS63773  
 ID AAS63773 standard; cDNA; 2984 BP.

XX AAS63773;  
 AC  
 XX

DT 29-JAN-2002 (first entry)  
 XX

XX Human prostate cDNA sequence #325.  
 DE

XX Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.  
 KW

XX Homo sapiens.  
 OS

XX WO2001.73032-A2.  
 PN

XX 04-OCT-2001.  
 PD

XX 27-MAR-2001; 2001WO-US09919.  
 XX

XX 27-MAR-2000; 2000US-0536857.  
 PR

XX 09-MAY-2000; 2000US-0568100.  
 PR

PR 12-MAY-2000; 2000US-0570737.  
 PR 13-JUN-2000; 2000US-0593793.  
 PR 27-JUN-2000; 2000US-0605783.  
 PR 10-AUG-2000; 2000US-0636215.  
 PR 29-AUG-2000; 2000US-0651236.  
 PR 06-SEP-2000; 2000US-0657279.  
 PR 02-OCT-2000; 2000US-0679426.  
 PR 10-OCT-2000; 2000US-0685166.  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 PI Fanger GR, Rector MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;  
 XX WPI; 2001-639232/73.  
 DR P-PSDB; AAU69771.  
 XX  
 DR New human prostate-specific polypeptides and polynucleotides useful for  
 PT the diagnosis and treatment of cancer, especially prostate cancer -  
 XX  
 PS Claim 1; Page 335-336; 579pp; English.  
 CC The invention relates to isolated prostate-specific  
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,  
 CC antibodies raised against the polypeptides (or antigenic epitopes  
 CC derived from them) and antigen-presenting cells expressing the  
 CC polypeptides. The antibodies are useful for detecting the presence of  
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and  
 CC the antigen-presenting cells are useful for stimulating and/or expanding  
 CC T cells specific for a tumour protein, and for inhibiting the development  
 CC of cancer, especially prostate cancer. Compositions comprising the  
 CC polynucleotide and/or polypeptide are useful for stimulating an immune  
 CC response, and for treating cancer. The oligonucleotide is useful for  
 CC detecting cancer. The present sequence is a prostate specific  
 CC polynucleotide of the invention.  
 XX

Sequence 2984 BP; 837 A; 659 C; 645 G; 842 T; 1 other;

Query Match 94.9%; Score 1836.4; DB 22; Length 2984;  
 Best Local Similarity 99.0%; Pred. No. 0;

Matches 1897; Conservative 0; Mismatches 12; Indels 7; Gaps 5;

QY 1 AATGGTATGCCAACTTAAGTATTTACAGGGTGGCCCAATAGACAGATGCACTGCTG 60  
 Db 1067 AATGGTATGCCAACTTAAGTATTTACAGGGTGGCCCAATAGACAGATGCACTGCTG 1126  
 QY 61 TGATTTAAGACAAAGCTGTATAACAGAACTCCACTGCAAGAGGGGGCCGGCCAGGA 120  
 Db 1127 TGATTTAAGACAAAGCTGTATAACAGAACTCCACTGCAAGAGGGGGCCGGCCAGGA 1186  
 QY 121 GAATCTCCGCTGTGCCAAGACAGGGCCTAAGAGGGTCTCCACACTGCTGTAGGGGCT 180  
 Db 1187 GAATCTCCGCTGTGCCAAGACAGGGCCTAAGAGGGTCTCCACACTGCTGTAGGGGCT 1246  
 QY 181 GTTGCAATTTTTATTAGTAGAAGTGGAAAGGCTCTTCTCAACTTTTTTCCCTTGGGC 240  
 Db 1247 GTTGCAATTTTTATTAGTAGAAGTGGAAAGGCTCTTCTCAACTTTTTTCCCTTGGGC 1306  
 QY 241 TGGAGAATTTAGAAATCAGAGTTTCTTGGAGTTTTCAGGCTATCATATATCTGTATCT 300  
 Db 1307 TGGAGAATTTAGAAATCAGAGTTTCTTGGAGTTTTCAGGCTATCATATATCTGTATCT 1366  
 QY 301 GAAAGGCACATAATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360  
 Db 1367 GAAAGGCACATAATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1426  
 QY 361 TTACTCACTAAGGGCTTCATTTTACTCCAGATTTTACTCTGGCTGCACCTAATTTATG 420  
 Db 1427 TTACTCACTAAGGGCTTCATTTTACTCCAGATTTTACTCTGGCTGCACCTAATTTATG 1486  
 QY 421 CCTCGCTTATTTAGCCCGAGATCTGCTCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 480



methods for prostate cancer. They can indicate the level of metastasis as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to AAH01318 represent polynucleotide and amino acid sequences used in the exemplification of the present invention.

[illegible]





Db	1901	AGCCTAGGCCCTGGGGAGCCCGAGTCCACTGAGCAAGCAAGGAGCTGAGTGAGCGCTTTTGC	1961
Qy	901	AGGAAAAAGGCTAAGAAAAAGGAAAAACCAATTCATAAAACAACAAGAAACTGTGCCAAATGC	960
Db	1961	AGGAAAAAGGCTAAGAAAAAGGAAAAACCAATTCATAAAACAACAAGAAACTGTGCCAAATGC	2020
Qy	961	TTTGGGAACTGTGTTATTATGCGCTATAATATGGGTCCCAAAAATGGGTAACTCTAGACTTCAGA	1020
Db	2021	TTTGGGAACTGTGTTATTATGCGCTATAATATGGGTCCCAAAAATGGGTAACTCTAGACTTCAGA	2080
Qy	1021	GAGAACTGAGCAGAGAGCAAAAGGAGAAATCTGGCTGTCTCTTCCATTTTCATCTCTGTTATCT	1080
Db	2081	GAGAACTGAGCAGAGAGCAAAAGGAGAAATCTGGCTGTCTCTTCCATTTTCATCTCTGTTATCT	2140
Qy	1081	CAGGTGAGCTGGTAGAGGGGAGACATTTAGAAAAAAATGAAAAACAACAACAACAATTACTAAT	1140
Db	2141	CAGGTGAGCTGGTAGAGGGGAGACATTTAGAAAAAAATGAAAAACAACAACAACAATTACTAAT	2200
Qy	1141	GAGGTACGCTGAGCGCTGGGAGTCTCTTGACTCACTACTTAATTCGGTTTAGTGAGAAA	1200
Db	2201	GAGGTACGCTGAGCGCTGGGAGTCTCTTGACTCACTACTTAATTCGGTTTAGTGAGAAA	2260
Qy	1201	CCTTTCAATTTCTTTTATTAGAGGGCCAGCTTACTTTGTTGGTGCAAAAATGCCAACAT	1260
Db	2261	CCTTTCAATTTCTTTTATTAGAGGGCCAGCTTACTTTGTTGGTGCAAAAATGCCAACAT	2320
Qy	1261	AAGTTAATAGAAAGTTGGCCAAATTTACCCCAATTTCTGTGGTTTGGGCTCCCAATTGCA	1320
Db	2321	AAGTTAATAGAAAGTTGGCCAAATTTACCCCAATTTCTGTGGTTTGGGCTCCCAATTGCA	2380
Qy	1321	ATGTTCAATGCGCAGCTGCTGTGACCGGAGGAGTACTAGCCGACACAAAAAGGCGAGG	1380
Db	2381	ATGTTCAATGCGCAGCTGCTGTGACCGGAGGAGTACTAGCCGACACAAAAAGGCGAGG	2440
Qy	1381	TAGCCTCAATTTGCTTCTGCTCTTTACATTTCTTTTAAATAAGCATTTAGTGTCTCAGTC	1440
Db	2441	TAGCCTCAATTTGCTTCTGCTCTTTACATTTCTTTTAAATAAGCATTTAGTGTCTCAGTC	2500
Qy	1441	CCTACTGAGTACTCTTTCTCTCCCTCTCTGAAATTTAAATCTTTTCAACTTGCATTTTGC	1500
Db	2501	CCTACTGAGTACTCTTTCTCTCCCTCTCTGAAATTTAAATCTTTTCAACTTGCATTTTGC	2560
Qy	1501	AAGGATTTACACATTTCACTGTGATGTATATGTGTGTCAGNGAAGAAAAAGTGTCTT	1560
Db	2561	AAGGATTTACACATTTCACTGTGATGTATATGTGTGTCAGNGAAGAAAAAGTGTCTT	2619
Qy	1561	TGTTTAAAAATTACTTGGTTTGTGAATCCATCTTGCTTTTTTCCCATTTGGAACCTAGTCATT	1620
Db	2620	TGTTTAAAAATTACTTGGTTTGTGAATCCATCTTGCTTTTTTCCCATTTGGAACCTAGTCATT	2679
Qy	1621	AACCCATCTCTGAACCTGGTAGAAAAACATCTGAAGAGCTAGTCTATCAGCATCTGCACGG	1680
Db	2680	AACCCATCTCTGAACCTGGTAGAAAAACATCTGAAGAGCTAGTCTATCAGCATCTGCACGG	2739
Qy	1681	TGAATTTGGATGGTTCTCAGAACCAATTTACCCGACAGAGCCTGTTCTATCTCTGTTTAAATA	1740
Db	2740	TGAATTTGGATGGTTCTCAGAACCAATTTACCCGACAGAGCCTGTTCTATCTCTGTTTAAATA	2799
Qy	1741	AATTAGTTTGGGTTCTCTACATGCATAACAAACCCCTGCTCCAATCTGTACATAAAAGTC	1800
Db	2800	AATTAGTTTGGGTTCTCTACATGCATAACAAACCCCTGCTCCAATCTGTACATAAAAGTC	2859
Qy	1801	TGTGACTTGAAGTTTAGTTCAGCACCCGCCCAAACTTTATTTTTCTATGTGTTTTTGTGCA	1860
Db	2860	TGTGACTTGAAGTTTAGTTCAGCACCCGCCCAAACTTTATTTTTCTATGTGTTTTTGTGCA	2919
Qy	1861	ACATATGAGTGTGTTTGAATAAAGTACCAATGTCTTTATTAATAAAAAAAAAAAAAA	1916
Db	2920	ACATATGAGTGTGTTTGAATAAAGTACCAATGTCTTTATTAATAAAAAAAAAAAAAA	2975

## RESULT 7



```

PR 02-OCT-2000; 2000US-0679426.
PR 10-OCT-2000; 2000US-0685166.
XX
PA (XUJ)/ XU J.
PA (DILL)/ DILLON D C.
PA (MITC)/ MITCHAM J L.
PA (HARL)/ HARLOCKER S L.
PA (JIAN)/ JIANG Y.
PA (KALO)/ KALOS M D.
PA (FANG)/ FANGER G R.
PA (RETT)/ RETTER M W.
PA (STOL)/ STOLK J A.
PA (DAYC)/ DAY C H.
PA (VEDV)/ VEDVICK T S.
PA (CART)/ CARTER D.
PA (LISX)/ LI S X.
PA (WANG)/ WANG A.
PA (SKEI)/ SKEIKY Y A W.
PA (HEPL)/ HEPLER W T.
PA (HEND)/ HENDERSON R A.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JH, Day CH, Vedvick TS,
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX
DR WPI; 2002-255649/30.
XX
PT New prostate-specific polynucleotides for diagnosing and treating
PT diseases, in particular prostate cancer, and as markers for the
PT progression of cancer
XX
PS Claim 1; SEQ ID NO(335); 47pp; English.
XX
CC The present invention provides prostate-specific coding sequences and
CC their encoded proteins. These can be used in the diagnosis and treatment
CC of cancers, particularly prostate cancer. The present sequence is a cDNA
CC described in the invention.
XX
SQ Sequence 2984 BP; 837 A; 659 C; 645 G; 842 T; 1 other;
XX
Query Match 94.9%; Score 1836.4; DB 24; Length 2984;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1897; Conservative 0; Mismatches 12; Indels 7; Gaps 5;
QY 1 AATGGTATGCCAACTTAAGTATTACAGGTGGCCCAATAGAACAGATGCACCTCGCTG 60
DB 1067 AATGGTATGCCAACTTAAGTATTACAGGTGGCCCAATAGAACAGATGCACCTCGCTG 1126
QY 61 TGATTTTAAGACAAAGCTGTATAACAGAACTCCACTGCAAGAGGGGGCCGGCCAGGA 120
DB 1127 TGATTTTAAGACAAAGCTGTATAACAGAACTCCACTGCAAGAGGGGGCCGGCCAGGA 1186
QY 121 GAATCTCCGCTTGCCAGACAGGGGCCCTAAGAGGGTCTCCACACTGCTGTAGGGGT 180
DB 1187 GAATCTCCGCTTGCCAGACAGGGGCCCTAAGAGGGTCTCCACACTGCTGTAGGGGT 1246
QY 181 GTTGCAATTTTATTAGTAGAAGTGGAAAGGCCCTCTCTCAACTTTTTTCCCTTGGGC 240
DB 1247 GTTGCAATTTTATTAGTAGAAGTGGAAAGGCCCTCTCTCAACTTTTTTCCCTTGGGC 1306
QY 241 TGGAGAAATTTAGAAATCAGAAATTTCTCGGAGTTTTCAGGCTATCATATATCTGTATCT 300
DB 1307 TGGAGAAATTTAGAAATCAGAAATTTCTCGGAGTTTTCAGGCTATCATATATCTGTATCT 1366
QY 301 GAAAGCAACATAATTTCTTCCCTCCCTTTTAAATTTTGGTTCCTTTTGCAGAA 360
DB 1367 GAAAGCAACATAATTTCTTCCCTCCCTTTTAAATTTTGGTTCCTTTTGCAGAA 1426
QY 361 TTACTCACTAAAGGGCTTCAATTTTGTAGTCCAGATTTTGTCTGGCTGCACCTTAATG 420
DB 1427 TTACTCACTAAAGGGCTTCAATTTTGTAGTCCAGATTTTGTCTGGCTGCACCTTAATG 1486
QY 421 CCTCGCTTATTAGCCCGAGATCTGGTCTTTTNTGNTTTTTTTTTTTTTCGCTCTCCC 480

```

```

1487 CCTCGCTTATTAGCCCGAGATCTGGTCTTTT---TTTTTTTTTTTTTTCGCTCTCC 1544
QY
481 CAAAGCTTTATCTGCTTGTGACCTTTTAAAAAGTTTGGGGGAGATTTCTGAATTTGGGTA 540
DB
1545 CAAAGCTTTATCTGCTTGTGACCTTTTAAAAAGTTTGGGGGAGATTTCTGAATTTGGCTA 1503
QY
541 AAAGACATGCATTTTAAAACTAGGCACTTCTTAATTTCTTCTTTTAAAAATACATAGC 600
DB
1604 AAAGACATGCATTTTAAAACTAG--CRACTCTTAATTTCTTCTTTAAAAATACATAGC 1661
QY
601 ATTAAATCCAAATCTTATTAAAGACCTGACAGCTTTGAAGAGTTCACCTACTGCAATTTAT 660
DB
1662 ATTAAATCCAAATCTTATTAAAGACCTGACAGCTTTGAAGAGTTCACCTACTGCAATTTAT 1721
QY
661 AGGACCTTCTGCTGTTCTGCTGTTTCAAGTCTGACAACTCTTCTGAGAAATCTTTGC 720
DB
1722 AGGACCTTCTGCTGTTCTGCTGTTTCAAGTCTGACAACTCTTCTGAGAAATCTTTGC 1781
QY
721 ATGACAGAGAGGTAAGAGGTATTGGATTTTACAGAGGAAGAACACAGCGCAGAAATGAAG 780
DB
1782 ATGACAGAGAGGTAAGAGGTATTGGATTTTACAGAGGAAGAACACAGCGCAGAAATGAAG 1841
QY
781 GGCAGGCTTACTGAGGCTGTCCAGTGGAGGCTCATGGGTGGGACATCGAAAAAGAGGC 840
DB
1842 GGCAGGCTTACTGA--GCTGTCCAGTGGAGGCTCATGGGTGGGACATCGAAAAAGAGGC 1900
QY
841 ASCCTAGGCCCTGGGGAGGCCAGTCCACCTGACGACAGAGGACTGAGTGAGGCTTTTGC 900
DB
1901 AGCTTAGGCCCTGGGGAGGCCAGTCCACCTGACGACAGAGGACTGAGTGAGGCTTTTGC 1960
QY
901 AGGAAAAGGCTAAGAAAAGGAAAACCAATTTAAAAACACAAAGAAACTGTCCTCAATGC 960
DB
1961 AGGAAAAGGCTAAGAAAAGGAAAACCAATTTAAAAACACAAAGAAACTGTCCTCAATGC 2020
QY
961 TTTGGAACTGTGTTTATTGCTTATAATGGGTCCCAAAATGGGTAACTAGACTTCAGA 1020
DB
2021 TTTGGAACTGTGTTTATTGCTTATAATGGGTCCCAAAATGGGTAACTAGACTTCAGA 2080
QY
1021 GAGATGACGACAGACGAAAGGAGAAATCTGCTGTCTTCTTCCATTTTCAATCTGTTATCT 1080
DB
2081 GAGATGACGACAGACGAAAGGAGAAATCTGCTGTCTTCTTCCATTTTCAATCTGTTATCT 2140
QY
1081 CAGGTGAGCTGTAGAGGGGAGACATTTAGAAAAAATAAACAACAAACAAATTAATAAT 1140
DB
2141 CAGGTGAGCTGTAGAGGGGAGACATTTAGAAAAAATAAACAACAAACAAATTAATAAT 2200
QY
1141 GAGGTACGCTGAGGCTGGGAGTCTCTGACTCCACTACTTAATTCCTGTTTGTAGGAAA 1200
DB
2201 GAGGTACGCTGAGGCTGGGAGTCTCTGACTCCACTACTTAATTCCTGTTTGTAGGAAA 2260
QY
1201 CCTTTCAATTTTCTTTTATTAGAGGGCCAGCTTACTGTTGGTGGCAAAATTTGCCAAT 1260
DB
2261 CCTTTCAATTTTCTTTTATTAGAGGGCCAGCTTACTGTTGGTGGCAAAATTTGCCAAT 2320
QY
1261 AAGTTAAATAGAAAGTTGGCCAAATTTCAACCCATTTTCTGTTGGTGGGCTCCACATTTGA 1320
DB
2321 AAGTTAAATAGAAAGTTGGCCAAATTTCAACCCATTTTCTGTTGGTGGGCTCCACATTTGA 2380
QY
1321 ATGTTCAATGCAAGCTGTGCTGACACCGACCGAGTACTAGCCAGCAACAAAGGCGAGG 1380
DB
2381 ATGTTCAATGCAAGCTGTGCTGACACCGACCGAGTACTAGCCAGCAACAAAGGCGAGG 2440
QY
1381 TAGCCTGAATTCCTTCTGCTCTTTTACATTTCTTTTAAATAAGCAATTTAGTGTCTCAGTC 1440
DB
2441 TAGCCTGAATTCCTTCTGCTCTTTTACATTTCTTTTAAATAAGCAATTTAGTGTCTCAGTC 2500
QY
1441 CCTACTGAGTACTCTTTCTCTCCCTCTCTCAATTTTAAATTTCTTCAACTTGCATTTGC 1500
DB
2501 CCTACTGAGTACTCTTTCTCTCCCTCTCTCAATTTTAAATTTCTTCAACTTGCATTTGC 2560
QY
1501 AAGGATTAACAAATTTCACTGTGATGATATTGTTGTGCAAGNAGAAAGAAAGTGTCTTT 1560

```

Db 2561 AAGATTACACATTTCACTGTGTATATTTGTTGCA-AAAAAAAAAAGTCTCTT 2619  
Qy 1561 TGTTTAAATTTACTTGTGTTGTAATCCATCTTCTTTTCCCATTTGGAATAGTCAAT 1620  
Db 2620 TGTTTAAATTTACTTGTGTTGTAATCCATCTTCTTTTCCCATTTGGAATAGTCAAT 2679  
Qy 1621 AACCATCTCTGAACCTGTGTAGAAAAACATCTGAAGAGCTAGTCTATCAGCATCTGACAGG 1680  
Db 2680 AACCATCTCTGAACCTGTGTAGAAAAACATCTGAAGAGCTAGTCTATCAGCATCTGACAGG 2739  
Qy 1681 TGAATTGGATGGTCTCAGAACCATTTCCACCCAGCAGCTGTTCTATCTGTTTAAATA 1740  
Db 2740 TGAATTGGATGGTCTCAGAACCATTTCCACCCAGCAGCTGTTCTATCTGTTTAAATA 2799  
Qy 1741 AATTAGTTTGGGTTCTCTACATGCATACAAACCCCTGCTCCCAATCTGCACATTAAGTC 1800  
Db 2800 AATTAGTTTGGGTTCTCTACATGCATACAAACCCCTGCTCCCAATCTGCACATTAAGTC 2859  
Qy 1801 TGTGACTTGAAGTTTGTAGTCAGCACCCCCACCAAACTTTATTTTCTATGTTTGTGCA 1860  
Db 2860 TGTGACTTGAAGTTTGTAGTCAGCACCCCCACCAAACTTTATTTTCTATGTTTGTGCA 2919  
Qy 1861 ACATATGAGTGTGTTTGAATAAAGTACCCATGCTGTTTATTAATAAATAAATAAATAA 1916  
Db 2920 ACATATGAGTGTGTTTGAATAAAGTACCCATGCTGTTTATTAATAAATAAATAAATAA 2975

RESULT 9  
ACA59581  
ID ACA59581 standard; cDNA; 2984 BP.  
AC ACA59581;  
XX  
DT 10-JUN-2003 (first entry)  
XX  
DE Prostate cancer therapy associated CDNA #324.  
XX  
KW Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;  
KW immunogen; cancer; prostate specific antigen; PSA;  
KW prostatic acid phosphatase; PAP; prostate specific membrane antigen;  
KW PSMA; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN US2002192763-A1.  
XX  
PD 19-DEC-2002.  
XX  
PF 29-JUN-2001; 2001US-0895793.  
XX  
PR 17-APR-2000; 2000US-157455P.  
PR 04-OCT-2000; 2000US-0679272.  
PR 28-MAR-2001; 2001US-0822827.  
XX  
PA (XUJ/) XU J.  
PA (DILL/) DILLON D C.  
PA (MITC/) MITCHAM J L.  
PA (HARL/) HARLOCKER S L.  
PA (JIAN/) JIANG Y.  
PA (KALO/) KALOS M D.  
PA (FANG/) FANGER G R.  
PA (RETT/) RETTER M W.  
PA (STOL/) STOLK J A.  
PA (DAYC/) DAY C H.  
PA (VEDV/) VEDVICK T S.  
PA (CART/) CARTER D.  
PA (LISX/) LI S X.  
PA (WANG/) WANG A.  
PA (SKEI/) SKEIKY Y A W.  
PA (HEPL/) HEPLER W T.  
PA (HEND/) HENDERSON R A.  
PA (HURA/) HURAL J.  
PA (MCNE/) MCNEILL P D.

(HOUG/) HOUGHTON R L.  
(DBAS/) Y DE BASSOLS C V.  
(FOYT/) FOY T M.  
Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;  
Mcneill PD, Houghton RL, Y De Bassols CV, Foy TM;  
WPI; 2003-352711/33.  
P-PSDB; ABU71662.  
New fusion protein comprising prostate-specific polypeptides, or its  
immunogenic portions, useful for diagnosing, preventing and/or treating  
cancer, particularly prostate cancer -  
Example 3; SEQ ID NO 335; 85pp; English.  
The invention describes a fusion protein comprising at least one amino  
acid sequence of immunogenic portions of any of the 3 sequences not  
defined in the specification, or sequences having at least 70 or 90 %  
sequence identity to any one of the 3 sequences defined in the USPTO  
web site, which is encoded by any of the 4 nucleotide sequences not  
defined in the specification. The fusion protein, composition and  
methods are useful for diagnosing, preventing and/or treating cancer,  
particularly prostate cancer. The proteins are useful as markers to  
indicate the presence or absence of cancer. This sequence  
represents a prostate cancer therapy associated cDNA.  
CC Note: The sequence data for this patent did not form part of the  
printed specification, but was obtained in electronic format directly  
from the US patent office at  
seqdata.uspto.gov/sequence.html?DocID=US20020192763.  
XX  
SQ. Sequence 2984 BP; 837 A; 659 C; 645 G; 842 T; 1 other;  
Query Match 94.9%; Score 1836.4; DB 25; Length 2984;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1897; Conservative 0; Mismatches 12; Indels 7; Gaps 5;  
Qy 1 AATGGTATGCCAACTTAAGTATTACAGGGTGGCCCAATAAGATGCACTCGCTG 60  
Db 1067 AATGGTATGCCAACTTAAGTATTACAGGGTGGCCCAATAAGATGCACTCGCTG 1126  
Qy 61 TGATTTTAAAGCAAGCTGTATAACACAACTCCACTGCAAGAGGNGGGCCGCCAGGA 120  
Db 1127 TGATTTTAAAGCAAGCTGTATAACAACTCCACTGCAAGAGGNGGGCCGCCAGGA 1186  
Qy 121 GAATCTCCGCTTGTCCAAGACAGGGGCTTAAGGAGGCTCTCCACACTGCTGCTAGGGCT 180  
Db 1187 GAATCTCCGCTTGTCCAAGACAGGGGCTTAAGGAGGCTCTCCACACTGCTGCTAGGGCT 1246  
Qy 181 GTTGCATTTTATTTAGTAGAAGTGGAAAGCCCTCTCTCAACTTTTTCCTTTGGGC 240  
Db 1247 GTTGCATTTTATTTAGTAGAAGTGGAAAGCCCTCTCTCAACTTTTTCCTTTGGGC 1306  
Qy 241 TGGAGATTTTAGAATCAGAAGTTTCTCTGAGTGTTCAGGCTATCATATATCTATCTCT 300  
Db 1307 TGGAGATTTTAGAATCAGAAGTTTCTCTGAGTGTTCAGGCTATCATATATCTATCTCT 1366  
Qy 301 GAAAGGCAACATAATTTCTTCCCTCCCTCTTCTTAAATTTTGTTTTCCTTTGACGAA 360  
Db 1367 GAAAGGCAACATAATTTCTTCCCTCCCTCTTCTTAAATTTTGTTTTCCTTTGACGAA 1426  
Qy 361 TTACTCACTAAAGGGCTTCATTTTAGTCCAGATTTTGTAGTGGCTGACCTATCTATG 420  
Db 1427 TTACTCACTAAAGGGCTTCATTTTAGTCCAGATTTTGTAGTGGCTGACCTATCTATG 1486  
Qy 421 CCTCGCTTATTTAGCCCGAGATCTGCTCTTTTNGTNTTTTTTTTTTCCGCTCC 480  
Db 1487 CCTCGCTTATTTAGCCCGAGATCTGCTCTTTT - TTTTTTTTTTCCGCTCC 1544  
Qy 481 CAAAGCTTTATCTGCTCTGACATTTTAAAGTGTGGGGCAGATCTGATTTGGGCTA 540

Db 1545 CAAAGCTTTATCTGTCTTGACTTTTAAAAAGTTTGGGGCAGATTCGAATT-GGCTA 1603  
Qy 541 AAAGACATGCAATTTTAAAACTAGGCAACTCTTAATTTCTTTTCTTTTAAAAATACATAGC 600  
Db 1604 AAAGACATGCAATTTTAAAACTAG--CAACTCTTTATTTCTTTCTTTTAAAAATACATAGC 1661  
Qy 601 ATTAATATCCCAATCTTAATTAAGACCTGACAGCTTGAGAGGTGACACTGCAATTTAT 660  
Db 1662 ATTAATATCCCAATCTTAATTAAGACCTGACAGCTTGAGAGGTGACACTGCAATTTAT 1721  
Qy 661 AGGACCTTCTGGTGGTCTCTGCTGTGTACGTTTGAAGTCTGACAACTCTTGAGAACTTTTGC 720  
Db 1722 AGGACCTTCTGGTGGTCTCTGCTGTGTACGTTTGAAGTCTGACAACTCTTGAGAACTTTGC 1781  
Qy 721 ATGACAGGAGGTGAAGAGGTATTTGATTTTACAGAGGAGAAACACAGCGCAGAAATGAAG 780  
Db 1782 ATGACAGGAGGTGAAGAGGTATTTGATTTTACAGAGGAGAAACACAGCGCAGAAATGAAG 1841  
Qy 781 GGCCAGGCTTACTGAGGCTGTCAGTGGAGGCTCATGGTGGGACATATGGAAAGAGGC 840  
Db 1842 GGCCAGGCTTACTGA-GCTGTCAGTGGAGGCTCATGGTGGGACATATGGAAAGAGGC 1900  
Qy 841 AGCTTAGGCCCTGGGAGGCCAGTCCACTGAGCAAGCAAGGGACTGAGTGAGCCTTTTTC 900  
Db 1901 AGCTTAGGCCCTGGGAGGCCAGTCCACTGAGCAAGCAAGGGACTGAGTGAGCCTTTTTC 1960  
Qy 901 AGGAAAGGCTAAGAAAGAGGAAAGCAATTTCTAAAAACAAACAGAAACTGTCCAAATGC 960  
Db 1961 AGGAAAGGCTAAGAAAGAGGAAAGCAATTTCTAAAAACAAACAGAAACTGTCCAAATGC 2020  
Qy 961 TTTGGGAACTGTTTATTTGCTTATTAATGGGTCCCAAAATGGGTAACTAGACTTCAGA 1020  
Db 2021 TTTGGGAACTGTTTATTTGCTTATTAATGGGTCCCAAAATGGGTAACTAGACTTCAGA 2080  
Qy 1021 GAGAATGACAGAGCAAGAGGAGAAATCTGGTCTCTTCCATTTTCAATTTCTGTTATCT 1080  
Db 2081 GAGAATGACAGAGCAAGAGGAGAAATCTGGTCTCTTCCATTTTCAATTTCTGTTATCT 2140  
Qy 1081 CAGGTGAGCTGTGAGAGGGAGAGCAATTAGAAAAAATGAAAAACAAACAAATTAAT 1140  
Db 2141 CAGGTGAGCTGTGAGAGGGAGAGCAATTAGAAAAAATGAAAAACAAACAAATTAAT 2200  
Qy 1141 GAGGTAGCTGAGGCTGGAGTCTCTGACTCCACTACTTAATTCGGTTTAGTGAGAA 1200  
Db 2201 GAGGTAGCTGAGGCTGGAGTCTCTGACTCCACTACTTAATTCGGTTTAGTGAGAA 2260  
Qy 1201 CCTTCAATTTCTTTTATTAAGAGGCGAGCTTACTGTTGGTGCAAAATTCGCAACAT 1260  
Db 2261 CCTTCAATTTCTTTTATTAAGAGGCGAGCTTACTGTTGGTGCAAAATTCGCAACAT 2320  
Qy 1261 AAGTTAATAGAAAGTTGGCCAAATTTCAACCCCAATTTTCTGTGGTTTGGGCTCCACATTCGA 1320  
Db 2321 AAGTTAATAGAAAGTTGGCCAAATTTCAACCCCAATTTTCTGTGGTTTGGGCTCCACATTCGA 2380  
Qy 1321 ATGTTCAATGCCAGTGTGCTGACACCGACGGAGTACTAGCCAGCAAAAGGCGAGG 1380  
Db 2381 ATGTTCAATGCCAGTGTGCTGACACCGACGGAGTACTAGCCAGCAAAAGGCGAGG 2440  
Qy 1381 TAGCCTGAATGCTTCTGCTCTTTTCAATTTCTTTTAAATAGCAATTTAGTCTCAGTC 1440  
Db 2441 TAGCCTGAATGCTTCTGCTCTTTTCAATTTCTTTTAAATAGCAATTTAGTCTCAGTC 2500  
Qy 1441 CCTACTGAGTACTCTTCTCTCCCTCTCTGAAATTAATTTCTTCACTTGAATTTTCG 1500  
Db 2501 CCTACTGAGTACTCTTCTCTCCCTCTCTGAAATTAATTTCTTCACTTGAATTTTCG 2560  
Qy 1501 AAGGATTACATTTCACTGTGATGATATGTTGTCAGNAGAAAGAAAAAGTGTCTT 1560  
Db 2561 AAGGATTACATTTCACTGTGATGATATGTTGTTGCA-AAAAAAGAAAAAGTGTCTT 2619  
Qy 1561 TGTTTAAATTAATCTGGTTGTGAATCCATCTGCTTTTCCCATGGAATGATTCAT 1620  
Db 2620 TGTTTAAATTAATCTGGTTGTGAATCCATCTGCTTTTCCCATGGAATGATTCAT 2679

Qy 1621 AACCCATCTCTGAACCTGGTAGAAAAACATCTGAAGAGCTAGTCTATCAGCATCTGCAGG 1680  
Db 2680 AACCCATCTCTGAACCTGGTAGAAAAACATCTGAAGAGCTAGTCTATCAGCATCTGCAGG 2739  
Qy 1681 TGAATGGATGGTCTCAGAACATTTACCCAGACAGCCTGTTCTATCTCTGTTTAAATA 1740  
Db 2740 TGAATGGATGGTCTCAGAACATTTACCCAGACAGCCTGTTCTATCTCTGTTTAAATA 2799  
Qy 1741 AATTAGTTTGGGTTCTCTACATGCATAACAAACCCCTGCTCCCAATCTGTCCATAAAAGTC 1800  
Db 2800 AATTAGTTTGGGTTCTCTACATGCATAACAAACCCCTGCTCCCAATCTGTCCATAAAAGTC 2859  
Qy 1801 TGTGACTTGAAGTTTAGTCAGCACCCCAACCAACTTTATTTTCTATCTGTTTTCGA 1860  
Db 2860 TGTGACTTGAAGTTTAGTCAGCACCCCAACCAACTTTATTTTCTATCTGTTTTCGA 2919  
Qy 1861 ACATATGAGTGTGTTGAAAAATAAGTACCATGTCTTTATTAANAANAANAANA 1916  
Db 2920 ACATATGAGTGTGTTGAAAAATAAGTACCATGTCTTTATTAANAANAANAANAANA 2975

RESULT 10  
ABV25139 standard; cDNA; 3227 BP.  
XX AC ABV25139;  
XX AC  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 25130.  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX Homo sapiens.  
XX WO200160860-A2.  
XX 23-AUG-2001.  
XX 20-FEB-2001; 2001WO-US05171.  
XX 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
PS Claim 1; Page 4902; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX Sequence 3227 BP; 879 A; 740 C; 765 G; 828 T; 15 other;  
SQ

Query Match  
Best Local Similarity 83.9%; Score 1624.6; DB 23; Length 3227;  
Matches 1752; Conservative 0; Mismatches 134; Indels 10; Gaps 4;

QY 1 AATGGTATGCCAATTAAGTATTTACAGGTGGCCCAATAGAACAAAGATGCACTCGCTG 60  
DB |||||  
QY 61 TGAATTTAAGACAAGCTGTATAAACAAGACTCCATGCAAGAGGNGGCGCGCCAGGA 120  
DB |||||  
QY 1382 TGAATTTAAGACAAGCTGTATAAACAAGACTCCATGCAAGAGGNGGCGCGCCAGGA 1441  
DB |||||  
QY 121 GAATCTCCGCTTGTCCAAGACAGGGGCTTAAGAGGGTCTCCACACTGCTGTAGGGCT 180  
DB |||||  
QY 1442 GAATCTCCGCTTGTCCAAGACAGGGGCTTAAGAGGGTCTCCACACTGCTGTAGGGCT 1501  
DB |||||  
QY 181 GTTGCAATTTTTTATTAGTAGAAGTGGNAAGGCTCTTCAACTTTTTTCCCTTGGGC 240  
DB |||||  
QY 1502 GTTGCAATTTTTTATTAGTAGAAGTGGNAAGGCTCTTCAACTTTTTTCCCTTGGGC 1561  
DB |||||  
QY 241 TGGAGATTTAGAACTCAGAAGTTTCTCGAGTTTTCAGGCTATCATATATCTGTATCT 1621  
DB |||||  
QY 1562 TGGAGATTTAGAACTCAGAAGTTTCTCGAGTTTTCAGGCTATCATATATCTGTATCT 1621  
DB |||||  
QY 301 GAAAGGCAACATAATCTTCTCCCTCTCTTTTAAATTTTGTCTCTTTTGCAGCA 360  
DB |||||  
QY 1622 GAAAGGCAACATAATCTTCTCCCTCTCTTTTAAATTTTGTCTCTTTTGCAGCA 1681  
DB |||||  
QY 361 TPACTCACTAAGGGCTTCATTTAGTCCAGATTTTGTAGTCTGCTGCTCACTAATATG 420  
DB |||||  
QY 1682 TPACTCACTAAGGGCTTCATTTAGTCCAGATTTTGTAGTCTGCTGCTCACTAATATG 1741  
DB |||||  
QY 421 CCTGCTTATTAGCCGAGATCTGGTCTTTTNTGNTTTTTTTTTTTTTTTTTTTCGCTCC 480  
DB |||||  
QY 1742 CCTGCTTATTAGCCGAGATCTGGTCTTTTNTGNTTTTTTTTTTTTTTTTTTTCGCTCC 1801  
DB |||||  
QY 481 CAAAGCTTTATCTGCTTGTACTTTTAAAGTTTGGGGCAGATTCGAATTTGGCTA 540  
DB |||||  
QY 1802 CAAAGCTTTATCTGCTTGTACTTTTAAAGTTTGGGGCAGATTCGAATTTGGCTA 1860  
DB |||||  
QY 541 AAGACATGCAATTTTAAACTAGGCACTTCTATTCTTTCTTTTAAATAACATAGC 600  
DB |||||  
QY 1861 AAGACATGCAATTTTAAACTAG--CAACTCTTATTCTTCTTTTAAATAACATAGC 1918  
DB |||||  
QY 601 ATTAATATCCCAATCTCTATTAAAGACCTGACAGCTTGAGAGTCACTACTGCAATTTAT 660  
DB |||||  
QY 1919 ATTAATATCCCAATCTCTATTAAAGACCTGACAGCTTGAGAGTCACTACTGCAATTTAT 1978  
DB |||||  
QY 661 AGGACCTTCTGGTGTCTGCTGTTTGAAGTCTGACAACTCTTGAGAACTTTTGC 720  
DB |||||  
QY 1979 AGGACCTTCTGGTGTCTGCTGTTTGAAGTCTGACAACTCTTGAGAACTTTTGC 2038  
DB |||||  
QY 721 ATGACAGAGGATTAAGAGGTATTGGAATTTTACAGAGGAGAACACAGCGCAGATGAAG 780  
DB |||||  
QY 2039 ATGACAGAGGATTAAGAGGTATTGGAATTTTACAGAGGAGAACACAGCGCAGATGAAG 2098  
DB |||||  
QY 781 GGCAGGCTTATGAGGCTGTCCAGTGGAGGCTCATGGTGGGACATGGAAAGAGAGGC 840  
DB |||||  
QY 2099 GGCAGGCTTATGGA--GCTGTCCAGTGGAGGCTCATGGTGGGACATGGAAAGAGAGGC 2157  
DB |||||  
QY 841 AGCTAGGCGCTGGGAGCCAGTCCACTGACAGAGGAGGAGCTGAGTGAAGCTTTTGC 900  
DB |||||  
QY 2158 AGCTAGGCGCTGGGAGCCAGTCCACTGACAGAGGAGGAGTGAAGTGAAGCTTTTGC 2217  
DB |||||  
QY 901 AGGAAAGGCTTAAGAAAAGGAAAAACCAATTTTAAACACAAAGAACTGTCCAAATGC 960  
DB |||||

RESULT 11  
ABV22893  
ID ABV22893 standard; cDNA; 3228 BP.  
XX AC  
XX ABV22893;  
XX

DB 2218 AGGAAAAGGCTAAGAAAAAGGAAAAACCATTTCTTAAAAACACAAAGAACTGTCCAAATGC 2277  
QY 961 TTTGGGAACCTGCTGTTTATTCCTTATTAATGGGTCCCAAAATGGGTAACTAGACTTCAGA 1020  
DB |||||  
QY 2278 TTTGGGAACCTGCTGTTTATTCCTTATTAATGGGTCCCAAAATGGGTAACTAGACTTCAGA 2337  
DB |||||  
QY 1021 GAGAATGAGCAGAGAGCAAGAGAGAAATCTGGCTGCTCCCTCCATTTTCAATCTGTTATCT 1080  
DB |||||  
QY 2338 GAGAATGAGCAGAGAGCAAGAGAGAAATCTGGCTGCTCCCTCCATTTTCAATCTGTTATCT 2397  
DB |||||  
QY 1081 CAGGTGAGCTGTGAGAGGAGAGCAATTAAGAAAAAATGAAAAAACAACAAATTAATAAT 1140  
DB |||||  
QY 2398 CAGGTGAGCTGTGAGAGGAGAGCAATTAAGAAAAAATGAAAAAACAACAAATTAATAAT 2457  
DB |||||  
QY 1141 GAGGTACGCTGAGGCGCTGGAGTCTCTTGAATCTTCACTTAAATTCGGTTAGTGAGAAA 1200  
DB |||||  
QY 2458 GAGGTACGCTGAGGCGCTGGAGTCTCTTGAATCTTCACTTAAATTCGGTTAGTGAGAAA 2517  
DB |||||  
QY 1201 CTTTCAATTTTCTTTTATTAGAAGGGCAGCTTACTGTTGGTGCAAAATTTGCCAACAT 1260  
DB |||||  
QY 2518 CTTTCAATTTTCTTTTATTAGAAGGGCAGCTTACTGTTGGTGCAAAATTTGCCAACAT 2577  
DB |||||  
QY 1261 AAGTTAAATAGAAGTTGGCCAAATTTACCCCAATTTTCTGCTGTTGGCTCCCAATTGCA 1320  
DB |||||  
QY 2578 AAGTTAAATAGAAGTTGGCCAAATTTACCCCAATTTTCTGCTGTTGGCTCCCAATTGCA 2637  
DB |||||  
QY 1321 ATGTTCAATGCGACGCTGCTGCTGACACCGACGGAGTACTAGCCAGACAAAAGCGAGG 1380  
DB |||||  
QY 2638 ATGTTCAATGCGACGCTGCTGCTGACACCGACGGAGTACTAGCCAGACAAAAGCGAGG 2697  
DB |||||  
QY 1381 TAGCTCGAATTTGCTTTCTGCTCTTTTAAATTAAGCAATTTAGTGTCTAGTC 1440  
DB |||||  
QY 2698 TAGCTCGAATTTGCTTTCTGCTCTTTTAAATTAAGCAATTTAGTGTCTAGTC 2757  
DB |||||  
QY 1441 CTTACTGAGTACTCTTTCTCTCCCTCTCTGTAATTTAAATTTCTTCACTTTGCAATTTGC 1500  
DB |||||  
QY 2758 CTTACTGAGTACTCTTTCTCTCCCTCTCTGTAATTTAAATTTCTTCACTTTGCAATTTGC 2817  
DB |||||  
QY 1501 AAGCAATACACATTTCACTGCTGATGATATATGTTGTCAGNGAAAAAGAAAAAGTG--- 1556  
DB |||||  
QY 2818 AAGCAATACACATTTCACTGCTGATGATATATGTTGTCAGNGAAAAAGAAAAAGTG--- 2877  
DB |||||  
QY 1557 --TCTTTGTTTAAATTTACTTTGTTGTAATCCATCTGCTTTTCCCATTTGGAACTA 1614  
DB |||||  
QY 2878 CTTTTTGTGAAAAAACCCTGGGTGGGAAACCCATTTGTTTTCCTTTTCCCATTTGGAACCT 2937  
DB |||||  
QY 1615 GTCAATTAACCATCTCTGAATCTGTAAGAAAAACATCTGAAGAGCTAGTCTATCAGCATCT 1674  
DB |||||  
QY 2938 GGTTTTAAACCCCTTTTAAACCGGAAAAAAGTGTGAAAAAGTTTTTCACCTTTT 2997  
DB |||||  
QY 1675 GACAGTGAATTTGGATGTTTCTCAGAACCAATTTTCAACCGAGAGAGCTGTTCTATCTCT 1734  
DB |||||  
QY 2998 AACGGGAAATGGTGGTTTAAACCCCTTTTCCCAACCGCGGGTTTTTATCCGCT 3057  
DB |||||  
QY 1735 TTAATTAATTTGTTGGTCTCTACATGCAATAACAAACCCCTGCTCCAATTTTCCCAAAA 1794  
DB |||||  
QY 3058 TAAAAAATAATTTGGGGTTTTTAAACAGGAAAAAACCCTGCTCCAATTTTCCCAAAA 3117  
DB |||||  
QY 1795 AAAGTCTGTGACTTGAAGTTTATGTCAGACCCCAACCACTTATTTTCTATGTTT 1854  
DB |||||  
QY 3118 AAATTTGGGAATTTGGAATTTAAATTTTCAACCCCAACCAATTTTAAATTTTCAAGGGTTT 3177  
DB |||||  
QY 1855 TTTGCAACATATGAGTGTGTTTGAATAAAGTATGCC 1890  
DB |||||  
QY 3178 TTTGGCAACAAAAAGGGTTTGAATAAAGTATGCC 3213  
DB |||||



Db 2638 ARGTTCAATGCCAGTCTGCTGCTAGCACCGACCGAGTACTAGCAGCACAAAAGGCGAGG 2697  
 QY 1381 TAGCCTGAATGCTTCTGCTCTTACATTTCTTTTAAATAAGCATTTAGTCTCAGTTC 1440  
 Db 2698 TAGCCTGAATGCTTCTGCTCTTACATTTCTTTTAAATAAGCATTTAGTCTCAGTTC 2757  
 QY 1441 CCTACTGAGTACTCTTCTCTCCCTCTCTGAACTTAATTTCTTCAACTTCCAAATTGC 1500  
 Db 2758 CCTACTGAGTACTCTTCTCTCCCTCTCTGAACTTAATTTCTTCAACTTCCAAATTGC 2817  
 QY 1501 AAGGATTACACATTTCACTGCTGATGATATTTGTTGCGAGNGAAGAAAAGTGG---- 1556  
 Db 2818 AAGGATTACACATTTCACTGCTGATGATATTTGTTGCGAGNGAAGAAAAGTGG---- 1556  
 QY 1557 --TCTTTGTTTAAATTTACTTGTGTTGTTGTAATCCATCTTCTTTTCCCATTTGAACTA 1614  
 Db 2878 CCTTTTGTGTAATAAAACCTGGTGGGNAACCCATTTGTTTCCCTTTGGAACCT 2937  
 QY 1615 GTCAATTAACCATCTCTGAAGTGTGTAAGAAACATCTGAAGAGCTAGTCTATCAGCATCT 1674  
 Db 2938 GGTTTTAACCCCTTTTAAACGGGGGAAAAAAACCTTGAAAAAGTGTTTTTCACCTTT 2997  
 QY 1675 GACAGTGAATTTGATGTTCTCAGAACCATTTTCCACGAGCAGCTGTTCTATCTGT 1734  
 Db 2998 TAACGGGGAAATGGTGGGTTTAAACCCCTTTTCCCAACGCGCGGTTTTATCCGG 3057  
 QY 1735 TTAATAAATAGTTTGGGTTCTCTACATGATACAAACCCCTGCTCCCAATCTGTACATA 1794  
 Db 3058 TTAATAAATAATTTGGGGTTTTTAAACAGGAAAAAAACCCCTGCTCCCAATTTGCCAAA 3117  
 QY 1795 AAGTCTGTGACTGAAGTTTGTAGTCAGACCCCTCCACCAACTTTATTTTCTATGTG--T 1852  
 Db 3118 AAAATTTGGAATTTGAATTTAATTTACCCCTCCCAAAATTTAATTTTCCAAGGGTT 3177  
 QY 1853 TTTTGTCAACATATGAGTGTGTTGTAATAAATAAGTACCC 1890  
 Db 3178 TTTTGTCAACAAAAAGGGTTTGAATAAATAAGCCCCC 3215

RESULT 12  
 ID ABV23209  
 XX ABV23209 standard; cDNA; 3228 BP.  
 AC ABV23209;  
 XX  
 DT 16-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 23200.  
 XX  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US05171.  
 XX  
 PR 17-FEB-2000; 2000US-183319P.  
 PR 16-MAR-2000; 2000US-189862P.  
 PR 25-MAY-2000; 2000US-207454P.  
 PR 09-JUN-2000; 2000US-211314P.  
 PR 18-JUL-2000; 2000US-219007P.  
 PR 13-DEC-2000; 2000US-255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Schlegel R, Endege WO, Monahan JE;  
 XX  
 DR WPI; 2001-662795/76.  
 XX

PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 XX for detecting presence of prostate cancer, stage of prostate cancer -  
 PS Claim 1; Page 4176; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for:  
 CC (a) assessing whether a patient is afflicted with prostate cancer;  
 CC (b) monitoring the progression of prostate cancer in a patient;  
 CC (c) assessing the efficacy of a test compound to inhibit prostate  
 CC cancer in a patient;  
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
 CC in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound;  
 CC (g) determining whether prostate cancer has metastasized in a patient;  
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
 CC patient;  
 CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 3228 BP; 879 A; 740 C; 765 G; 829 T; 15 other;  
 SQ

Query Match 83.7%; Score 1620.6; DB 23; Length 3228;  
 Best Local Similarity 92.6%; Pred. No. 0;  
 Matches 1757; Conservative 0; Mismatches 129; Indels 12; Gaps 5;  
 QY 1 AATGTTATGCCAACTTAAGTATTTACAGGGTGGCCCAATAAGACAAGATGCTCGCTG 60  
 Db 1322 AATGTTATGCCAACTTAAGTATTTACAGGGTGGCCCAATAAGACAAGATGCTCGCTG 1381  
 QY 61 TGATTTAAGCAAGCTGTATATAACAGAACTCCACTGCAAGAGGGGGCCGCGCAGGA 120  
 Db 1382 TGATTTAAGCAAGCTGTATATAACAGAACTCCACTGCAAGAGGGGGCCGCGCAGGA 1441  
 QY 121 GAATCTCGCTTGTCCAAAGACAGGGGCTTAAGAGGGTCTCCACTGCTGTAGGGGCT 180  
 Db 1442 GAATCTCGCTTGTCCAAAGACAGGGGCTTAAGAGGGTCTCCACTGCTGTAGGGGCT 1501  
 QY 181 GTTGCAATTTTATTTAGTAAAGTGAAGAGGCTCTTCTCAACTTTTTCCTTGGGC 240  
 Db 1502 GTTGCAATTTTATTTAGTAAAGTGAAGAGGCTCTTCTCAACTTTTTCCTTGGGC 1561  
 QY 241 TGGAGAATTTAGAAATCAGAGTTTCTGGAGTTTTCAGGCTATCATATATCTATCT 300  
 Db 1562 TGGAGAATTTAGAAATCAGAGTTTCTGGAGTTTTCAGGCTATCATATATCTATCT 1621  
 QY 301 GAAAGGCAACATTAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 360  
 Db 1622 GAAAGGCAACATTAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1681  
 QY 361 TTAAGTCAATTAAGGGCTTCAATTTTAGTCAAGATTTTATTTAGTCTGGCTTAATG 420  
 Db 1682 TTAAGTCAATTAAGGGCTTCAATTTTAGTCAAGATTTTATTTAGTCTGGCTTAATG 1741  
 QY 421 CTTCTGCTTATTTAGCCGAGATCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 480  
 Db 1742 CTTCTGCTTATTTAGCCGAGATCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1801  
 QY 481 CAAAGCTTTATCTGCTTCTGACTTTTAAAGTTTGGGGCAGATTTCTGAATTTGGGCTA 540  
 Db 1802 CAAAGCTTTATCTGCTTCTGACTTTTAAAGTTTGGGGCAGATTTCTGAATTTGGGCTA 1860  
 QY 541 AAAGACATGCAATTTTAAAGCTAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 600  
 Db 1861 AAAGACATGCAATTTTAAAGCTAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1918  
 QY 601 ATTAATCCCAATCTTATTTAAGAGCTGACAGCTTTGAGAAGTCTACTCTGATTTAT 660  
 Db 1919 ATTAATCCCAATCTTATTTAAGAGCTGACAGCTTTGAGAAGTCTACTCTGATTTAT 1978  
 QY 661 AGGACCTTCTGCTGGTCTGCTGTTTACCTTTGAGTCTGACAACTCTTGGAAATCTTTG 720



QY 1 AATGTTATGCAACTTAAGTATTATTAAGGCTGGCCCAATAGAACAGATGCACTCGCTG 60  
Db 1322 AATGTTATGCAACTTAAGTATTATTAAGGCTGGCCCAATAGAACAGATGCACTCGCTG 1381  
QY 61 TGATTTAAAGACAGCTGTATAACAGAACTCACTGCAAGAGGCGGCGGCGCAGGA 120  
Db 1382 TGATTTAAAGACAGCTGTATAACAGAACTCACTGCAAGAGGCGGCGGCGCAGGA 1441  
QY 121 GAATCTCCGCTTGTCCAAAGACAGGGGCTTAAGAGAGGCTCTCCACACTGCTGCTAGGGCT 180  
Db 1442 GAATCTCCGCTTGTCCAAAGACAGGGGCTTAAGAGAGGCTCTCCACACTGCTGCTAGGGCT 1501  
QY 181 GTTGCAATTTTATTTAGTAGAAGTGAAGGCTCTCTCAACTTTTTTCCCTTGGC 240  
Db 1502 GTTGCAATTTTATTTAGTAGAAGTGAAGGCTCTCTCAACTTTTTTCCCTTGGC 1561  
QY 241 TGGAGAAATTTAGAAATCAGAAATTTCTCGAGTTTTCAGGCTATCATATATCTGATCCT 300  
Db 1562 TGGAGAAATTTAGAAATCAGAAATTTCTCGAGTTTTCAGGCTATCATATATCTGATCCT 1621  
QY 301 GAAAGCAACATTAATTTCTCTCCCTCTCTTTTAAATTTTGTGTTCTTTTGGCAGCA 360  
Db 1622 GAAAGCAACATTAATTTCTCTCCCTCTCTTTTAAATTTTGTGTTCTTTTGGCAGCA 1681  
QY 361 TTACTCACTAAAGGCTCTCATTTTGTAGTCCAGATTTTGTGTTCTTTTGGCAGCA 420  
Db 1682 TTACTCACTAAAGGCTCTCATTTTGTAGTCCAGATTTTGTGTTCTTTTGGCAGCA 1741  
QY 421 CCTCGTTATTTAGCCGAGATCTGGTCTTTTTTNTGTTTNTTTTTTTTTTTCGCTCTCC 480  
Db 1742 CCTCGTTATTTAGCCGAGATCTGGTCTTTTTTTTTTTTTTTTTTTTTTTCGCTCTCC 1801  
QY 481 CAAAGCTTTATCTGCTTCACTTTTAAAGTTTGGGCGCAGATCTGAATTTGGGCTA 540  
Db 1802 CAAAGCTTTATCTGCTTCACTTTTAAAGTTTGGGCGCAGATCTGAATTTGGGCTA 1860  
QY 541 AAAGACATGCAATTTTAAAGCTAGGCACTTTTATTTTCTTTTAAATTTACATAGC 600  
Db 1861 AAAGACATGCAATTTTAAAGCTAGGCACTTTTATTTTCTTTTAAATTTACATAGC 1918  
QY 601 ATTAATCCCAATCCCTATTTAAAGCTGAGCTGAGAGGCTCACTCACTCAATAT 660  
Db 1919 ATTAATCCCAATCCCTATTTAAAGCTGAGCTGAGAGGCTCACTCACTCAATAT 1978  
QY 661 AGGACCTTCTGGTGGTCTGCTGTTAGCTGTTGAAGCTGACATCTGAGATCTTTGC 720  
Db 1979 AGGACCTTCTGGTGGTCTGCTGTTAGCTGTTGAAGCTGACATCTGAGATCTTTGC 2038  
QY 721 ATGCAGAGGAGTAAAGGATTTGGATTTTCAAGAGGAGAAACAGCGCAGATGAAG 780  
Db 2039 ATGCAGAGGAGTAAAGGATTTGGATTTTCAAGAGGAGAAACAGCGCAGATGAAG 2098  
QY 781 GGCCAGGCTTACTGAGGCTGTCCAGTGGAGGCTCATGGTGGGACATGGAAAGAGGC 840  
Db 2099 GGCCAGGCTTACTGA-GCTGTCAAGTGGAGGCTCATGGTGGGACATGGAAAGAGGC 2157  
QY 841 AGCTAGGCTTGGGAGGCCAGTCCACTGAGCAAGAGGACTGAGTGAGCCTTTTGC 900  
Db 2158 AGCTAGGCTTGGGAGGCCAGTCCACTGAGCAAGAGGACTGAGTGAGCCTTTTGC 2217  
QY 901 AGGAAAGGCTAAGAAAAGGAAACCACTTAAACACAAAGAACTGTCCAAATGC 960  
Db 2218 AGGAAAGGCTAAGAAAAGGAAACCACTTAAACACAAAGAACTGTCCAAATGC 2277  
QY 961 TTTGGGAACGTGTTTATTTGCTTATTAATGGTCCCAAAATGGGTAACTAGACTTCA 1020  
Db 2278 TTTGGGAACGTGTTTATTTGCTTATTAATGGTCCCAAAATGGGTAACTAGACTTCA 2337  
QY 1021 GAGAATGACAGAGCAAGAGAAATCTGCTGCTCTTCAATTTTCAATCTGTTATCT 1080  
Db 2338 GAGAATGACAGAGCAAGAGAAATCTGCTGCTCTTCAATTTTCAATCTGTTATCT 2397  
QY 1081 CAGGTGAGCTGTGAGGGGAGACATTAGAAAAAATGAAACAAACAAACAACTTACTAAT 1140

Db 2398 CAGGTGAGCTGTGAGGGGAGACATTAGAAAAAATGAAACAAACAACTTACTAAT 2457  
QY 1141 GAGTAGCTGAGCCCTGGAGTCTCTTGACTCCACTACTTAATTCGTTTGTGAGAA 1200  
Db 2458 GAGTAGCTGAGCCCTGGAGTCTCTTGACTCCACTACTTAATTCGTTTGTGAGAA 2517  
QY 1201 CCTTTCAAATTTCTTTTATTAGAAAGGCGCAGCTTACTGTTGGTGGCAAAATTCGCAACAT 1260  
Db 2518 CCTTTCAAATTTCTTTTATTAGAAAGGCGCAGCTTACTGTTGGTGGCAAAATTCGCAACAT 2577  
QY 1261 AAGTTAATAGAAAGTTGGCAATTTTCAACCCCAATTTTCTGCTTTTGGGCTCCACATTGCA 1320  
Db 2578 AAGTTAATAGAAAGTTGGCAATTTTCAACCCCAATTTTCTGCTTTTGGGCTCCACATTGCA 2637  
QY 1321 ATGTTCAATGCGCAGCTGCTGACACCGGAGTACTAGCCAGCAACAAAGCAGGG 1380  
Db 2638 ATGTTCAATGCGCAGCTGCTGACACCGGAGTACTAGCCAGCAACAAAGCAGGG 2697  
QY 1381 TAGCTGAATTTGCTTCTGCTCTTTTACATTTCTTTTAAATTAAGCATTTAGTGTCACTC 1440  
Db 2698 TAGCTGAATTTGCTTCTGCTCTTTTACATTTCTTTTAAATTAAGCATTTAGTGTCACTC 2757  
QY 1441 CCTACTGAGTACTCTTCTCTCCCTCTCTGAAATTTTAAATTTTCACTTTGCAATTTGC 1500  
Db 2758 CCTACTGAGTACTCTTCTCTCCCTCTCTGAAATTTTAAATTTTCACTTTGCAATTTGC 2817  
QY 1501 AAGGATTAACATTTCACTGTGATGTATATTTGCTGAGNGAAAGAAAGAGTG- 1556  
Db 2818 AAGGATTAACATTTCACTGTGATGTATATTTGCTGAGNGAAAGAAAGAGTG- 2877  
QY 1557 -TCTTTCTTAAATTTACTTGGTTTGTGAATTCATCTGCTTTTCCCAATTTGGAACCTA 1614  
Db 2878 CTTTCTTTTAAATTTACTTGGTTTGTGAATTCATCTGCTTTTCCCAATTTGGAACCTA 2937  
QY 1615 GTCAATTAACCCATCTCTGAACTGCTAGAAACAACTCTGAAGAGCTAGTCTATCAGCATCT 1674  
Db 2938 GGTTTTAAACCCCTTTTAAAGGGGAAACAACTTTGAAAGAGTTTTCACACCTT 2997  
QY 1675 GACAGGTGAATTTGATGGTCTCAGAACCAATTTCAACAGACAGCTGTTTCTATCTCT 1734  
Db 2998 TACCGGGAATTTGGGTTTAAACCCCTTTTCCCAACAGCGGGTTTTCATCCG 3057  
QY 1735 TTAATAAATTTAGTTTGGGTTCTCTACATGATGAACCAACCTGCTTCAATCTGTACATA 1794  
Db 3058 TTAATAAATTTAGTTTGGGTTTAAACCCCTTTTCCCAACAGCGGGTTTTCATCCG 3117  
QY 1795 AAGTCTGTGACTTGAATTTAGTTCAGACCCCAACCACTTTATTTTCTATGTG--T 1852  
Db 3118 AAAATTTGGGAATTTGGAATTTAATTCACCCCAACCACTTTAATTTTCCAGGGTT 3177  
QY 1853 TTTTTCACACATATGATGTTTGAATTAATTAAGTACCC 1890  
Db 3178 TTTTTCACACAAAGGGTTTGAATTAATTAAGTACCC 3215

RESULT 14  
ABV23456  
ID ABV23456 standard; cDNA; 3228 BP.  
XX  
AC ABV23456;  
XX  
XX 16-SEP-2002 (first entry)  
XX Human prostate expression marker cdna 23447.  
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX pharmacogenomic marker; gene; ss.  
XX Homo sapiens.  
XX  
PN WO200160860-A2.  
XX





Search completed: August 26, 2003, 14:33:31  
Job time : 649 secs

```
QY 841 AGCTAGGCCCTGGGAGCCAGTCCACTGAGCAAGGAGTCTGAGTGAGCCTTTTGC 900
Db 2158 AGCTAGGCCCTGGGAGCCAGTCCACTGAGCAAGGAGTCTGAGTGAGCCTTTTGC 2217
QY 901 AGGAAAGGCTAAGAAAAGGAAAACCATTTCTAAACACAAAGAAACTGTCCAAATGC 960
Db 2218 AGGAAAGGCTAAGAAAAGGAAAACCATTTCTAAACACAAAGAAACTGTCCAAATGC 2277
QY 961 TTGGGAACCTGTCTTTATTGSCCTATATATGGGTCCCAAAATGGGTAAACCTAGACTTCAGA 1020
Db 2278 TTGGGAACCTGTCTTTATTGSCCTATATATGGGTCCCAAAATGGGTAAACCTAGACTTCAGA 2337
QY 1021 GAGAAATGAGCAGAGCAAGAGAAATCTGGCTGTCTTCCATTTTCATTTCTGTTATCT 1080
Db 2338 GAGAAATGAGCAGAGCAAGAGAAATCTGGCTGTCTTCCATTTTCATTTCTGTTATCT 2397
QY 1081 CAGGTGAGCTGTAGAGGGGAGACATTTAGAAAAAATGAAACAAACAAATTAATAAT 1140
Db 2398 CAGGTGAGCTGTAGAGGGGAGACATTTAGAAAAAATGAAACAAACAAATTAATAAT 2457
QY 1141 GAGGTACGCTGAGGCCCTGGGAGTCTCTTGACTCCACTACTTAATCCGTTTAGTGAGAAA 1200
Db 2458 GAGGTACGCTGAGGCCCTGGGAGTCTCTTGACTCCACTACTTAATCCGTTTAGTGAGAAA 2517
QY 1201 CCTTTCAAATTTCTTTTATTAGAGGGCCAGCTTACTGTGGTGGCAAAATGGCCAACT 1260
Db 2518 CCTTTCAAATTTCTTTTATTAGAGGGCCAGCTTACTGTGGTGGCAAAATGGCCAACT 2577
QY 1261 AAGTTAATAGAAAGTTGGCCAAATTTCCACCCCAATTTCTGTGGTTGGGCTCCACATTGCA 1320
Db 2578 AAGTTAATAGAAAGTTGGCCAAATTTCCACCCCAATTTCTGTGGTTGGGCTCCACATTGCA 2637
QY 1321 ATGTTCAATGCCACGTGCTGACACCGAGTACTAGCCAGCACAAGGCGAGG 1380
Db 2638 ATGTTCAATGCCACGTGCTGACACCGAGTACTAGCCAGCACAAGGCGAGG 2697
QY 1381 TAGCCTGAATTTGCTGCTCTTTTAAATTAAGCAATTTAGTGTCTCAGTC 1440
Db 2698 TAGCCTGAATTTGCTGCTCTTTTAAATTAAGCAATTTAGTGTCTCAGTC 2757
QY 1441 CCTACTGAGTACTCTCTCCCTCTCTGAAATTAATTTCTTCAACTTGCATTTGC 1500
Db 2758 CCTACTGAGTACTCTCTCCCTCTCTGAAATTAATTTCTTCAACTTGCATTTGC 2817
QY 1501 AAGGATTACACATTTCACTGTGATGTATATTTGTTGCAAGGAAAAGAAAGTG- --- 1556
Db 2818 AAGGATTACACATTTCACTGTGATGTATATTTGTTGCAAGGAAAAGAAAGGCGC 2877
QY 1557 --TCCTTTGTTTAAATTAATTTGTTGTAATCCATCTTGTCTTTTCCCATTTGGAACATA 1614
Db 2878 CCTTTTTCGTAAGAAAAAAGCTGGTGGGAAACCCATTTGGTTTTTCCCTTTGGAACCT 2937
QY 1615 GTCATTAACCCATCTCTGAACCTGGTAGAAAACATCTGAAGAGCTAGTCTATCAGCATCT 1674
Db 2938 GGTTTTAAACCCCTTTTAAACGGGAAAAAAGCTTTGAAAAAGTTTTTTTTCACCTTT 2997
QY 1675 GACAGGTGAATTTGATGTTCTCAGAACCATTTTCAACCCAGACAGCCTGTTTCTATCCTGT 1734
Db 2998 TAACGGGGAATTTGGTGGGTTTAAACCCCTTTTCCCAACGCGGGTTTTTATCCGG 3057
QY 1735 TTAATAAATTTAGTTGGGTTCTCTACATGCAATAACAAACCTGCTCCAACTGTGCACATA 1794
Db 3058 TTAATAAATAATTTGGGGTTTTTAACAGGAAAAAAGCCCTGGTCCAAATTTGGCCAAAA 3117
QY 1795 AAAGTCTGTGACTTGAAGTTTAGTCAGACACCCCAACCACTTTATTTTCTATGTG--T 1852
Db 3118 AAAATTTGGGAATTTGAATTTAATTCACCCCAACCACTTTTATTTTCCAAAGGGTT 3177
QY 1853 TTTTTCACATATAGTGTGTTTGAATAAAGTATCCC 1890
Db 3178 TTTTTCACATATAGTGTGTTTGAATAAAGTATCCC 3215
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2003, 14:22:42 ; Search time 173 Seconds  
(without alignments)  
4939.409 Million cell updates/sec

Title: US-09-700-700-1  
Perfect score: 1936  
Sequence: 1 aatggtatgccaacttaagt.....gggcggccgcgactagtga 1936

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/2/ina/5A COMB.seq.\*
- 2: /cgn2\_6/prodata/2/ina/5B COMB.seq.\*
- 3: /cgn2\_6/prodata/2/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/2/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/2/ina/PCUTUS COMB.seq.\*
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1836.4	94.9	2984	4	US-09-439-313-335
2	1836.4	94.9	2984	4	US-09-352-616A-141
3	1836.4	94.9	2984	4	US-09-232-149A-141
4	1388.8	71.7	2051	4	US-09-328-475C-334
5	641.2	33.1	1024	4	US-09-328-475C-81
6	628.4	32.5	1024	4	US-09-328-475C-96
7	628.2	32.4	1024	4	US-09-328-475C-82
8	622	32.1	1024	4	US-09-328-475C-95
9	615.6	31.8	1013	4	US-09-328-475C-14
10	584.6	30.2	948	4	US-09-328-475C-23
11	582.2	30.1	720	4	US-09-328-475C-319
12	581	30.0	744	4	US-09-328-475C-318
13	539.6	27.9	980	4	US-09-328-475C-19
14	469.4	24.2	484	4	US-09-439-313-434
15	469.4	24.2	484	4	US-09-352-616A-434
16	421.6	21.8	1020	4	US-09-328-475C-102
17	421.6	21.8	1021	4	US-09-328-475C-103
18	411	21.2	822	4	US-09-328-475C-223
19	378.2	19.5	497	4	US-09-328-475C-222
20	350.4	18.1	366	3	US-09-020-956-115
21	350.4	18.1	366	3	US-09-030-607-115
22	350.4	18.1	366	4	US-09-439-313-115
23	350.4	18.1	366	4	US-09-352-616A-115
24	350.4	18.1	366	4	US-09-232-149A-115
25	332	17.1	335	3	US-09-020-956-141
26	332	17.1	335	3	US-09-030-607-141
27	332	17.1	335	4	US-09-439-313-141

C 28	332	17.1	335	4	US-09-352-616A-141	Sequence 141, App
C 29	332	17.1	335	4	US-09-232-149A-141	Sequence 141, App
C 30	332	16.7	323	3	US-09-020-956-128	Sequence 128, App
C 31	332	16.7	323	3	US-09-030-607-128	Sequence 128, App
C 32	332	16.7	323	3	US-09-439-313-128	Sequence 128, App
C 33	332	16.7	323	4	US-09-352-616A-128	Sequence 128, App
C 34	332	16.7	323	4	US-09-232-149A-128	Sequence 128, App
C 35	319.4	16.5	499	4	US-09-328-475C-185	Sequence 185, App
36	308.6	15.9	317	4	US-09-439-313-453	Sequence 453, App
37	308.6	15.9	317	4	US-09-352-616A-453	Sequence 453, App
38	288	14.9	305	4	US-09-439-313-295	Sequence 295, App
39	288	14.9	305	4	US-09-352-616A-295	Sequence 295, App
40	288	14.9	305	4	US-09-232-149A-295	Sequence 295, App
C 41	282.4	14.6	624	4	US-09-439-313-443	Sequence 443, App
C 42	282.4	14.6	624	4	US-09-352-616A-443	Sequence 443, App
C 43	268.8	13.9	332	3	US-09-030-607-207	Sequence 207, App
44	268.8	13.9	332	4	US-09-439-313-207	Sequence 207, App
45	268.8	13.9	332	4	US-09-352-616A-207	Sequence 207, App

ALIGNMENTS

RESULT 1

US-09-439-313-335  
; Sequence 335, Application US/09439313  
; Patent No. 6329505  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan Louise  
; APPLICANT: Jiang Yuqi  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Kalos, Michael  
; APPLICANT: Fanger, Gary  
; APPLICANT: Retter, Mark  
; APPLICANT: Solk, John  
; APPLICANT: Day, Craig  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
; FILE REFERENCE: 210121.427C9  
; CURRENT APPLICATION NUMBER: US/09/439,313  
; CURRENT FILING DATE: 1999-11-12  
; NUMBER OF SEQ ID NOS: 575  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 335  
; LENGTH: 2984  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-439-313-335

Query Match	94.9%;	Score	1836.4;	DB	4;	Length	2984;
Best Local Similarity	99.0%;	Pred. No.	0;				
Mismatches	1897;	Conservative	0;	Mismatches	12;	Indels	7;
Gaps	5;						
Qy	1	AATGTTATGCCAACTTAAGTATTTCAGGTCGCCCAATAAGAACAGATGACATCGCTG	60				
Db	1067	AATGTTATGCCAACTTAAGTATTTCAGGTCGCCCAATAAGAACAGATGACATCGCTG	1126				
Qy	61	TGATTTTAAGACAAGCTGTATAAAGAGAACTTCCTCAAGAGGNGGCGCGGCCAGGA	120				
Db	1127	TGATTTTAAGACAAGCTGTATAAAGAGAACTTCCTCAAGAGGNGGCGCGGCCAGGA	1186				
Qy	121	GAATCTCGCTTGTCCAAAGACAGGGGCTTAAGAGGGTCTCCACACTGCTGTAGGGGT	180				
Db	1187	GAATCTCGCTTGTCCAAAGACAGGGGCTTAAGAGGGTCTCCACACTGCTGTAGGGGT	1246				
Qy	181	GTTCATTTTATTATTAGTAGAAGCTGGAAGAGGCTCTTCTCAACTTTTTCCTTTGGGC	240				
Db	1247	GTTCATTTTATTATTAGTAGAAGCTGGAAGAGGCTCTTCTCAACTTTTTCCTTTGGGC	1306				
Qy	241	TGGAGATTTAGAAATCAGAAGTTTCTCGAGTTTTCAGGCTATCATATATACTGTATCT	300				





QY 1 AATGGTATGCCAATTAAGTATTTACAGGTGGCCCAATAGAAAGATGCACTGGCTG 60  
Db |||||  
1067 AATGGTATGCCAATTAAGTATTTACAGGTGGCCCAATAGAAAGATGCACTGGCTG 1126  
QY 61 TGATTTTAAAGCAAGCTGTATAAACAAGAACTCCACTGCAAGAGGNGGCGCGGCGAGA 120  
Db |||||  
1127 TGATTTTAAAGCAAGCTGTATAAACAAGAACTCCACTGCAAGAGGNGGCGCGGCGAGA 1186  
QY 121 GAATCTCCGCTTCTCAAGACAGGGGCTTAAGAGGGTCTCCACTGCTGTAGGGCT 180  
Db |||||  
1187 GAATCTCCGCTTCTCAAGACAGGGGCTTAAGAGGGTCTCCACTGCTGTAGGGCT 1246  
QY 181 GTTGCAATTTTATTTAGTAGAAAGTGAAAGGCTCTCTCAACTTTTCCCTTGGC 240  
Db |||||  
1247 GTTGCAATTTTATTTAGTAGAAAGTGAAAGGCTCTCTCAACTTTTCCCTTGGC 1306  
QY 241 TGGAGATTTAGTAATCAGAGTTTCTGGAGTTTTCAGGCTATCATATATCTGTATCCT 300  
Db |||||  
1307 TGGAGATTTAGTAATCAGAGTTTCTGGAGTTTTCAGGCTATCATATATCTGTATCCT 1366  
QY 301 GAAAGCAACATAATTTCTTCTTCCCTCTCTTTTAAATTTTGTGTTCTTTTGGAGCA 360  
Db |||||  
1367 GAAAGCAACATAATTTCTTCTTCCCTCTCTTTTAAATTTTGTGTTCTTTTGGAGCA 1426  
QY 361 TTACTCATAAAGGCTTCANTTTAGTCCAGATTTTATAGTCTGGCTGCACCTAATTTATG 420  
Db |||||  
1427 TTACTCATAAAGGCTTCANTTTAGTCCAGATTTTATAGTCTGGCTGCACCTAATTTATG 1486  
QY 421 CCTCGCTTATTTAGCCGAGATCTGGCTCTTTTNTGNTTTTNTTNTTNTTNTTNTTNTT 480  
Db |||||  
1487 CCTCGCTTATTTAGCCGAGATCTGGCTCTTTTNTGNTTTTNTTNTTNTTNTTNTTNTT 1544  
QY 481 CAAAGCTTATCTGTCTTACTTTTAAAGTTTGGGGCAGATTTCTGAATTTGGGCTA 540  
Db |||||  
1545 CAAAGCTTATCTGTCTTACTTTTAAAGTTTGGGGCAGATTTCTGAATTTGGGCTA 1603  
QY 541 AAAGCATGCAATTTTAAAGTAGCAACTTTTATTTCTTCTTTTAAAGTTTGGGCTA 600  
Db |||||  
1604 AAAGCATGCAATTTTAAAGTAGCAACTTTTATTTCTTCTTTTAAAGTTTGGGCTA 1661  
QY 601 ATTAATCCCAATCTTATTTAAAGCTGACAGCTTGAGAGGTGACACTGCAATTTAT 1721  
Db |||||  
1662 ATTAATCCCAATCTTATTTAAAGCTGACAGCTTGAGAGGTGACACTGCAATTTAT 1721  
QY 661 AGGACCTTCTGGTGTCTGCTGTAGTGAAGTCTGCAATCTTGAAGTCTTTGC 720  
Db |||||  
1722 AGGACCTTCTGGTGTCTGCTGTAGTGAAGTCTGCAATCTTGAAGTCTTTGC 1781  
QY 721 ATGCAGAGGAGTAAAGAGTATTTGATTTTACAGAGGAGACACAGCGCAGATGAAG 780  
Db |||||  
1782 ATGCAGAGGAGTAAAGAGTATTTGATTTTACAGAGGAGACACAGCGCAGATGAAG 1841  
QY 781 GGCAGGCTTACTGAGCTGTCCAGTGGAGGCTCATGGGTGGGACATGGAAGAGAGG 840  
Db |||||  
1842 GGCAGGCTTACTGA-CTGTCCAGTGGAGGCTCATGGGTGGGACATGGAAGAGAGG 1900  
QY 841 AGGCTAGGCTTGGAGGCTTCCACTGACAGCAAGGAGTGAAGTGGCTTTTGC 900  
Db |||||  
1901 AGGCTAGGCTTGGAGGCTTCCACTGACAGCAAGGAGTGAAGTGGCTTTTGC 1960  
QY 901 AGGAAAGGCTAAGAAAGGAAACCAATTTTAAACACACACAGAACTGTCGAATGC 960  
Db |||||  
1961 AGGAAAGGCTAAGAAAGGAAACCAATTTTAAACACACACAGAACTGTCGAATGC 2020  
QY 961 TTTGGGAAGTGTATTTGCTTATTTAGTGGTCCCAAAATGGGTAACTGAGTTCAGA 1020  
Db |||||  
2021 TTTGGGAAGTGTATTTGCTTATTTAGTGGTCCCAAAATGGGTAACTGAGTTCAGA 2080  
QY 1021 GAGATGACAGAGCAAGAGGAGAACTGCTGTCTTCCATTTTCAATTTCTGTATCT 1080  
Db |||||  
2081 GAGATGACAGAGCAAGAGGAGAACTGCTGTCTTCCATTTTCAATTTCTGTATCT 2140

QY 1081 CAGGTAGCTGTAGAGGGGAGACATTTAGAAATAATGAAACAAACAATTAATAAT 1140  
Db |||||  
2141 CAGGTAGCTGTAGAGGGGAGACATTTAGAAATAATGAAACAAACAATTAATAAT 2200  
QY 1141 GAGTACGCTGAGGCCCTGGGAGTCTCTTGACTCCACTACTTAATTCGGTTAGTGAGAA 1200  
Db |||||  
2201 GAGTACGCTGAGGCCCTGGGAGTCTCTTGACTCCACTACTTAATTCGGTTAGTGAGAA 2260  
QY 1201 CTTTCAATTTCTTTTATTTAGAGGGCCAGCTTACTGTGTGGGCAAAATGGCCAAAT 1260  
Db |||||  
2261 CTTTCAATTTCTTTTATTTAGAGGGCCAGCTTACTGTGTGGGCAAAATGGCCAAAT 2320  
QY 1261 AGTTTAAATAGAAAGTTGGCCAAATTTCAACCCCAATTTCTGTTGGTGGGCAAAATGGCCAAAT 1320  
Db |||||  
2321 AGTTTAAATAGAAAGTTGGCCAAATTTCAACCCCAATTTCTGTTGGTGGGCAAAATGGCCAAAT 2380  
QY 1321 ATGTTCAATGCCACGCTGCTGCTGACACCGAGCGAGTACTAGCCAGCAAAAGGCGAGG 1380  
Db |||||  
2381 ATGTTCAATGCCACGCTGCTGCTGACACCGAGCGAGTACTAGCCAGCAAAAGGCGAGG 2440  
QY 1381 TAGCCTGAATTCGTTTCTGCTCTCTTACATTTCTTTTAAATTAAGCATTTAGTGTCAAGT 1440  
Db |||||  
2441 TAGCCTGAATTCGTTTCTGCTCTCTTACATTTCTTTTAAATTAAGCATTTAGTGTCAAGT 2500  
QY 1441 CCTACTGAGTACTTTTCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500  
Db |||||  
2501 CCTACTGAGTACTTTTCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2560  
QY 1501 AAGGATTAACATTTTCACTGTGATGATATTTGTTGTCAGNGAAGAAAGAGTCTT 1560  
Db |||||  
2561 AAGGATTAACATTTTCACTGTGATGATATTTGTTGTCAGNGAAGAAAGAGTCTT 2619  
QY 1561 TGTTTAAATTAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1620  
Db |||||  
2620 TGTTTAAATTAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2679  
QY 1621 AACCCATCTGTAAGTGTGTAAGAAACATCTGAAGAGTGTCTATCAGCATCTGACAGG 1680  
Db |||||  
2680 AACCCATCTGTAAGTGTGTAAGAAACATCTGAAGAGTGTCTATCAGCATCTGACAGG 2739  
QY 1681 TGAATTTGATGTTTCTCAGAACCAATTTTCAACCCAGACAGCTGTTCTCTCTCTCTCTCT 1740  
Db |||||  
2740 TGAATTTGATGTTTCTCAGAACCAATTTTCAACCCAGACAGCTGTTCTCTCTCTCTCTCT 2799  
QY 1741 AATTTAGTTTGGTCTCTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
Db |||||  
2800 AATTTAGTTTGGTCTCTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2859  
QY 1801 TGTGACTTGAAGTTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1860  
Db |||||  
2860 TGTGACTTGAAGTTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2919  
QY 1861 ACATATGAGTGTGTTGAAATAAGTACCATGCTCTTTTATTTAAATAAATAAATAAATAA 1916  
Db |||||  
2920 ACATATGAGTGTGTTGAAATAAGTACCATGCTCTTTTATTTAGAAATAAATAAATAAATAA 2975

## RESULT 4

US-09-328-475C-334/c  
; Sequence 334, Application US/09328475C  
; Patent No. 6476207  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Jimmy  
; APPLICANT: Astel, Jon H.  
; APPLICANT: Carroll III, Eddie  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Ford, Donna M.  
; APPLICANT: Monahan, John E.  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Steinmann, Kathleen E.  
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT  
; TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER  
; FILE REFERENCE: 1532.002/200130.463

Qy	1289	CCCAATTTCTGTGGTTTGGGCTCCACATTTGCAATGTTTCAATGCGCATGTGCTGCTGACACC	1348
Db	978	CCCAATTTCTGTGGTTTGGGCTCCACATTTGCAATGTTTCAATGCGCATGTGCTGCTGACACC	919
Qy	1349	GACCGAGTACTAGCCAGCAGCAAAAGCGAGGTAGCTGAAATGCTTTCTGCTCTTTTACA	1408
Db	918	GACCGAGTACTAGCCAGCAGCAAAAGCGAGGTAGCTGAAATGCTTTCTGCTCTTTTACA	859
Qy	1409	TTTCTTTTAAAAATAAGCATTTAGTGCTCAGTCCCTACTCTGAGTACTCTTTCTCTCCCTCC	1468
Db	858	TTTCTTTTAAAAATAAGCATTTAGTGCTCAGTCCCTACTCTGAGTACTCTTTCTCTCCCTCC	799
Qy	1469	TCTGAAATTTAAATTTCTTTCAACTTTGCAATTTGCAAGGATTTACACATTTCACTGTGATGTAT	1528
Db	798	TCTGAAATTTAAATTTCTTTCAACTTTGCAATTTGCAAGGATTTACACATTTCACTGTGATGTAT	739
Qy	1529	ATTGTGTGTCAGNGAGAAAAAGAGTGTCTTTGTTTAAAAATTTACTTTGGTTTCTGGAATCC	1588
Db	738	ATTGTGTGTCAGNGAGAAAAAGAGTGTCTTTGTTTAAAAATTTACTTTGGTTTCTGGAATCC	680
Qy	1589	ATCTTGCTTTTCCCAATTTGGAACCTAGTCAATTAACCCATCTCTGAACTGTTGAGAAAAACA	1648
Db	679	ATCTTGCTTTTCCCAATTTGGAACCTAGTCAATTAACCCATCTCTGAACTGTTGAGAAAAACA	620
Qy	1649	TCTGAAGAGCTAGTCTATCAGCATCTGACAGGTGAATTTGCAATGTTCTCAGAACCATTTTC	1708
Db	619	TCTGAAGAGCTAGTCTATCAGCATCTGACAGGTGAATTTGCAATGTTCTCAGAACCATTTTC	560
Qy	1709	ACCAGACAGCCTGTTTCTATCTCTGTTTAAATAAATTTAGTTTGGGTTTCTCTACATGCATAA	1768
Db	559	ACCAGACAGCCTGTTTCTATCTCTGTTTAAATAAATTTAGTTTGGGTTTCTCTACATGCATAA	500
Qy	1769	CAAAACCTGCTCCAAATCTGTGCACATATAAAGTCTGTGACCTTGAAGTTAGTCAGACACCCC	1828
Db	499	CAAAACCTGCTCCAAATCTGTGCACATATAAAGTCTGTGACCTTGAAGTTAGTCAGACACCCC	440
Qy	1829	ACCAAACTTTATTTTCTATGTTTCTGTTTGGCAACATATGAGTGTGTTTGAATAAAGTAC	1888
Db	439	ACCAAACTTTATTTTCTATGTTTCTGTTTGGCAACATATGAGTGTGTTTGAATAAAGTAC	380
Qy	1889	CCATGCTTTTATTAAA	1904
Db	379	CCATGCTTTTATTAGA	364

RESULT 5

US-09-328-475C-81/c

; Sequence 81, Application US/09328475C

; Patent No. 6476207

; GENERAL INFORMATION:

; APPLICANT: Zhang, Jimmy

; APPLICANT: Astel, Jon H.

; APPLICANT: Carroll III, Eddie

; APPLICANT: Endege, Wilson O.

; APPLICANT: Ford, Donna M.

; APPLICANT: Monahan, John E.

; APPLICANT: Schlegel, Robert

; APPLICANT: Steinmann, Kathleen E.

; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT

; TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER

; FILE REFERENCE: 1532, 002/200130, 463

; CURRENT APPLICATION NUMBER: US/09/328,475C

; CURRENT FILING DATE: 1999-06-09

; NUMBER OF SEQ ID NOS: 341

; SOFTWARE: FASTSEQ for Windows Version 3.0

; SEQ ID NO 81

; LENGTH: 1024

; TYPE: DNA

; ORGANISM: Homo Sapien

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1)...(1024)

; OTHER INFORMATION: n = A, T, C or G  
US-09-328-475C-81

Query Match	33.1%	Score 641.2	DB 4	Length 1024
Best Local Similarity	97.3%	Pred. No. 3.8e-168		
Matches 681	Conservative 0	Mismatches 16	Indels 3	Gaps 3
448	CTTTTNTGNTNTTTTTTTTTTTTTTTCGCTCTCCCAAGCTTTATCTGCTTGACATTTTA	507		
799	CTCTTTNNNNNTTTTTTTTTTTCNGTCTCTCCCAAGCTTTATCTGCTTGACATTTNA	740		
508	AAAAAGTTGGGGCAGATCTGAAATGGGCTAAAGACATGCATTTTAAACTAGGCA	567		
739	AAAAAGTTGGGGCAGATCTGAAAT -GGTTAAAGACATGCATTTTAAACTA -GCA	682		
568	ACTTCTTATTTCTTTCTTTTAAATAATCATGCATTAATCCCAATCTTATTTAAAGAC	627		
681	ACTTCTTATTTCTTTCTTTTAAATAATCATGCATTAATCCCAATCTTATTTAAAGAC	622		
628	CTGACAGCTTGAGAAGCTCACTACTGCAATTTATAGGACCTTCGGTGGTTCTGCTGTAC	687		
621	CTGACAGCTTGAGAAGCTCACTACTGCAATTTATAGGACCTTCGGTGGTTCTGCTGTAC	562		
688	GTTTGAAGTCTGACAAATCTTGAGATCTTTTGATGTCAGAGAGGTAAGAGGTATTTGAT	747		
561	GTTTGAAGTCTGACAAATCTTGAGATCTTTTGATGTCAGAGAGGTAAGAGGTATTTGAT	502		
748	TTTTCAGAGGAAGAACACAGGCAGAAATGAAGGGCCAGGCTTACTGAGGCTGCCAGTG	807		
501	TTTTCAGAGGAAGAACACAGGCAGAAATGAAGGGCCAGGCTTACTGAGGCTGCCAGTG	443		
808	GAGGCTCATGGGTGGGACATGGAAGAGAGGCGAGCTTAGGCCCTGGGAGCCCAAGTCCA	867		
442	GAGGCTCATGGGTGGGACATGGAAGAGAGGCGAGCTTAGGCCCTGGGAGCCCAAGTCCA	383		
868	CTGACAGCAAGGCACTGAGTGAGCCTTTTGAGGAAAGGCTTAAGAAAAGGAAACC	927		
382	CTGACAGCAAGGCACTGAGTGAGCCTTTTGAGGAAAGGCTTAAGAAAAGGAAACC	323		
928	ATTTCTAAACAACAACAGAACTGTCCAAATCTTTGGGAATGTTTATGTCCTATAA	987		
322	ATTTCTAAACAACAACAGAACTGTCCAAATCTTTGGGAATGTTTATGTCCTATAA	263		
988	TGGGTCCCCAAATGGGTAACCTAGACTTCAGAGAAATGACGACAGCAAGAGGAAA	1047		
262	TGGGTCCCCAAATGGGTAACCTAGACTTCAGAGAAATGACGACAGCAAGAGGAAA	203		
1048	TCTGGCTGTCTTCCATTTTCATTCTGTTATCTCAGGTGAGCTGGTAGGGGAGACATT	1107		
202	TCTGGCTGTCTTCCATTTTCATTCTGTTATCTCAGGTGAGCTGGTAGGGGAGACATT	143		
1108	AGAAAAAATGAACAACAACAACAATTAATGAGGTAC	1147		
142	AGAAAAAATGAACAACAACAACAATTAATGAGGTAC	103		

## RESULT 6

US-09-328-475C-96/c  
 ; Sequence 96, Application US/09328475C  
 ; Patent No. 6476207  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Zhang, Jimmy  
 ; APPLICANT: Astel, Jon H.  
 ; APPLICANT: Carroll III, Eddie  
 ; APPLICANT: Endege, Wilson O.  
 ; APPLICANT: Ford, Donna M.  
 ; APPLICANT: Monahan, John E.  
 ; APPLICANT: Schlegel, Robert  
 ; APPLICANT: Steitmann, Kathleen E.  
 ; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT  
 ; TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER  
 ; FILE REFERENCE: 1532.002/200130.463  
 ; CURRENT APPLICATION NUMBER: US/09/328.475C

```

; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 96
; LENGTH: 1024

```

Query Match	32.5%;	Score 628.4;	DB 4;	Length 1024;
Best Local Similarity	97.4%;	Pred. No. 1.3e-164;		
Matches 678;	Conservative	0;	Mismatches 13;	Indels 5; Gaps 4;
QY	453	TTNTGNTTTTTTTTTTTTTTTTTTTTTTCGTCCTCCCAAGCTTTATCTGT-CITGACTTTTTAAAAA	511	
Db	795	TTTTNNNNTTTTTTTTTTTCNGTCTCCCAAGCTTTATCGTCTTGACCTTTTAAAAA	736	
QY	512	AGTTTGGGGCAGATTCTGAATTTGGGCTTAAAGACATGCAATTTTAAACTAGCACTT	571	
Db	735	AGTTTGGGGCAGATTCTGAATT-GGTTAAAGACATGCAATTTTAAAACTAG--CAATT	679	
QY	572	CTTATTCTCTTTCCTTTTAAAAATACATAGCAATTAATCCCAAACTCTATTTAAAGACCTGA	631	
Db	678	CTTATTCTCTTTCCTTTTAAAAATACATAGCAATTAATCCCAAACTCTATTTAAAGACCTGA	619	
QY	632	CAGCTTCAGAAAGTCACTACTGCAATTTATAGGACCTTCTGGTGGTCTGCTGTACGTTT	691	
Db	618	CAGCTTCAGAAAGTCACTACTGCAATTTATAGGACCTTCTGGTGGTCTGCTGTACGTTT	559	
QY	692	GAAGCTCGACAACTCTTGAGAACTTTTGCAATGACAGAGGTAAGAGGTATTGGATTTTC	751	
Db	558	GAAGCTCGACAACTCTTGAGAACTTTTGCAATGACAGAGGTAAGAGGTATTGGATTTTC	499	
QY	752	ACAGAGGAGAAACACAGCGGAGAAATGAAGGGCCAGCGCTTACTGAGCGTGTCCAGTGGAGG	811	
Db	498	ACAGAGGAGAAACACAGCGGAGAAATGAAGGGCCAGCGCTTACTGA-CTGTGTCAGTGGAGG	440	
QY	812	GCTCATGGTGGGACATGGAAGAGAGCGACTAGGCCCTGGGGAGGCCCACTGA	871	
Db	439	GCTCATGGTGGGACATGGAAGAGAGCGACTAGGCCCTGGGGAGGCCCACTGA	380	
QY	872	GCAAGCAGGGACTGAGTGAGCGCTTTTCAGGAAAAAGGCTAAGAAAAAGGAAACCACTC	931	
Db	379	GCAAGCAGGGACTGAGTGAGCGCTTTTCAGGAAAAAGGCTAAGAAAAAGGAAACCACTC	320	
QY	932	TAAAAACAACAAGAAACTGTCCAAATGCTTTGGGAACTGTGTTTATGTCTATAATGGG	991	
Db	319	TAAAAACAACAAGAAACTGTCCAAATGCTTTGGGAACTGTGTTTATGTCTATAATGGG	260	
QY	992	TCCCCAAAATGGGTAACTAGACTTCAGAGAAATGACGAGAGCAAAAGGAAATCTG	1051	
Db	259	TCCCCAAAATGGGTAACTAGACTTCAGAGAAATGACGAGAGCAAAAGGAAATCTG	200	
QY	1052	GCTGCTCTTCCATTTTCATTTCTGTATTCTCAGGTGAGCTGGTAGAGGGGAGACATTAGAA	1111	
Db	199	GCTGCTCTTCCATTTTCATTTCTGTATTCTCAGGTGAGCTGGTAGAGGGGAGACATTAGAA	140	
QY	1112	AAAAATGAACACAAACAACTTACTTAATGAGGTAC	1147	
Db	139	AAAAATGAACACAAACAACTTACTTAATGAGGTAC	104	

## RESULT 7

US-09-328-475C-82  
; Sequence 82, Application US/09328475C  
; Patent No. 6476207  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Jimmy  
; APPLICANT: Astel, Jon H.

APPLICANT: Carroll III, Eddie  
APPLICANT: Endege, Wilson O.  
APPLICANT: Ford, Donna M.  
APPLICANT: Monahan, John E.  
APPLICANT: Schlegel, Robert  
APPLICANT: Steinmann, Kathleen E.  
TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT  
ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER  
FILE REFERENCE: 1532.002/200130.463  
CURRENT APPLICATION NUMBER: US/09/328.475C  
CURRENT FILING DATE: 1999-06-09  
NUMBER OF SEQ ID NOS: 341  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 82  
LENGTH: 1024  
TYPE: DNA  
ORGANISM: Homo Sapien  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)..(1024)  
OTHER INFORMATION: n = A,T,C or G  
US-09-328-475C-82

Query Match 32.4%; Score 628.2; DB 4; Length 1024;  
Best Local Similarity 94.9%; Pred. No. 1.5e-164;  
Matches 690; Conservative 0; Mismatches 32; Indels 5; Gaps 4;

QY 422 CTCGCTATTATAGCCGAGATCTGGTCTTTTNTGNTTTTTTTTTTTTTCCTGCTCCCC 481  
DB 75 CGCCCTTAGCGTGTGCGGCCGAGGTACTCTTTTTTTTTTTTTTTTTTTCCTGCTCCCC 134

QY 482 AAAGCTTTATCTGCTGTGACTTTTTAAAAAGTTGGGGCAGATCTGAATTTGGCTAA 541  
DB 135 AAAGCTTTATCTGCTGTGACTTTTTAAAAAGTTGGGGCAGATCTGAATTTGGCTAA 193

QY 542 AAGACATGATTTTAAAACTAGCACTTCTTATTCTTCTTTTCTTTTAAAAATCATAGCA 601  
DB 194 AAGACATGATTTTAAAACTAGCACTTCTTATTCTTCTTTTCTTTTAAAAATCATAGCA 251

QY 602 TTAATCCCAATCCCTATTAAAGACCTGACGCTTGAGAGGTCACTACTGCAATTTATA 661  
DB 252 TTAATCCCAATCCCTATTAAAGACCTGACGCTTGAGAGGTCACTACTGCAATTTATA 311

QY 662 GGACCTTCTGGTGTCTGCTGTTAGCTTTGAAGTCTGCAATCTTGAAGTCTTTTGA 721  
DB 312 GGACCTTCTGGTGTCTGCTGTTAGCTTTGAAGTCTGCAATCTTGAAGTCTTTTGA 371

QY 722 TGACAGGAGGTAAAGGTATTGGATTTTCAAGAGAGCAACACAGCGCAGATGAAGG 781  
DB 372 TGACAGGAGGTAAAGGTATTGGATTTTCAAGAGAGCAACACAGCGCAGATGAAGG 431

QY 782 GCCAGGCTTACTGAGGCTGTCCAGTGGAGGCTCATGGTGGGACATGGAAGAAGGCA 841  
DB 432 GCCAGGCTTACTGA-GCTGTCCAGTGGAGGCTCATGGTGGGACATGGAAGAAGGCA 490

QY 842 GCCTAGGCCCTGGGGAGCCAGTCCACTGAGCAAGCAAGGCACTGAGTGGAGCCCTTTTGA 901  
DB 491 GCCTAGGCCCTGGGGAGCCAGTCCACTGAGCAAGCAAGGCACTGAGTGGAGCCCTTTTGA 550

QY 902 GGAAGGCTTAAGAAAAGGAAACCATTTCTAAACACACACAGAACTGTCCAAATGCT 961  
DB 551 GGAAGGCTTAAGAAAAGGAAACCATTTCTAAACACACACAGAACTGTCCAAATGCT 610

QY 962 TTGGGAACCTGTGTTTATGCTTATATGGTCCCAAAATGGGTAACTTAGACTTCAGAG 1021  
DB 611 TTGGGAACCTGTGTTTATGCTTATATGGTCCCAAAATGGGTAACTTAGACTTCAGAG 670

QY 1022 AGAATGAGCAGAGCAAGGAGAAATCTGGCTGTCTCTTCCATTTTCACTTCTGTTATCTC 1081  
DB 671 AGAATGAGCAGAGCAAGGAGAAATCTGGCTGTCTCTTCCATTTTCACTTCTGTTATCTC 730

QY 1082 AGGTGAGCTGTGAGGGAGACATTTAGAAAAAATGAAACAA-CAAAACAATTAATA 1140  
DB 1140

Db 731 AGGTGAACGTGTAAGAGGAGAGACATTTGAAAAAATGAAACNACCAAAACCATTACTAAT 790

QY 1141 GAGGTAC 1147  
DB 791 GAGGTAC 797

RESULT 8  
US-09-328-475C-95  
Sequence 95, Application US/09328475C  
Patent No. 6476207  
GENERAL INFORMATION:  
APPLICANT: Zhang, Jimmy  
APPLICANT: Astel, Jon H.  
APPLICANT: Carroll III, Eddie  
APPLICANT: Endege, Wilson O.  
APPLICANT: Ford, Donna M.  
APPLICANT: Monahan, John E.  
APPLICANT: Schlegel, Robert  
APPLICANT: Steinmann, Kathleen E.  
TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT  
ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER  
FILE REFERENCE: 1532.002/200130.463  
CURRENT APPLICATION NUMBER: US/09/328.475C  
NUMBER OF SEQ ID NOS: 341  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 95  
LENGTH: 1024  
TYPE: DNA  
ORGANISM: Homo Sapien  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)..(1024)  
OTHER INFORMATION: n = A,T,C or G  
US-09-328-475C-95

Query Match 32.1%; Score 622; DB 4; Length 1024;  
Best Local Similarity 95.0%; Pred. No. 8e-163;  
Matches 690; Conservative 0; Mismatches 30; Indels 6; Gaps 5;

QY 453 TTNTGNTTTTTTTTTTTTTCCTGCTCCCCAAAGCTTTATCTGCTGTGACTTTTAAAAA 512  
DB 101 TTTTTTTTTTTTTTTTTTTTCCTGCTCCCCAAAGCTTTATCTGCTGTGACTTTTAAAAA 160

QY 513 GTTTCGGGCGAGTCTGCAATTTGGCTAAAGACATGATTTTAAACTAGCACTTC 572  
DB 161 GTTTCGGGCGAGTCTGCAATTTGGCTAAAGACATGATTTTAAACTAG--CAACTC 217

QY 573 TTATTTCTTTTCTTTTAAAAATACATAGCATTTAAATCCCAATCCCTATTAAAGACCTGAC 632  
DB 218 TTATTTCTTTTCTTTTAAAAATACATAGCATTTAAATCCCAATCCCTATTAAAGACCTGAC 277

QY 633 AGCTTGAGAGGTCTACTGCAATTTATAGACCTTCTGGTGGTCTGCTGTGACTGTTG 692  
DB 278 AGCTTGAGAGGTCTACTGCAATTTATAGGACCTTCTGGTGGTCTGCTGTGACTGTTG 337

QY 693 AGCTTGCAATCTTGGAGATCTTTGATGAGAGGTAAAGAGGTATTTGATTTTCA 752  
DB 338 AGCTTGCAATCTTGGAGATCTTTGATGAGAGGTAAAGAGGTATTTGATTTTCA 397

QY 753 CAGAGGAAGAACACAGCCGAGAAATGAAGGGCCAGGCTTACTGAGGCTGTCCAGTGGAGG 812  
DB 398 CAGAGGAAGAACACAGCCGAGAAATGAAGGGCCAGGCTTACTGAG--CTGTCCAGTGGAGG 456

QY 813 CTCATGGTGGACATGAAAGAGGAGGAGGCTAGGCTTGGGGAGCCAGTCCACTGAG 872  
DB 457 CTCATGGTGGACATGAAAGAGGAGGAGGCTAGGCTTGGGGAGCCAGTCCACTGAG 516

QY 873 CAAGCAAGGGACTGAGTGGAGCTTTTGGAGGAAGGCTTAAGAAAAAGGAAACCACTCT 932  
DB 517 CAAGCAAGGGACTGAGTGGAGCTTTTGGAGGAAGGCTTAAGAAAAAGGAAACCACTCT 576



QY 508 AAAAAAGTTGGGGCAGATTCTGAAATGGGCTAAAGACATGCAATTTTAAACCTAGGCA 567  
Db |||||||  
QY 62 AAAAAAGTTGGGGCAGATTCTGAAAT-GGCTAAAGACATGCAATTTTAAACCTAG-CC 118  
Db |||||||  
QY 568 ACTCTTATTTCTTTCTTTAAATAATACATAGCAATTAATCCCAATCTTATTTAAAGAC 627  
Db |||||||  
QY 119 AACTCTTATTTCTTTCTTTAAATAATACATAGCAATTAATCCCAATCTTATTTAAAGAC 178  
Db |||||||  
QY 628 CTGACAGCTTTGAGAAGTCTACTGCAATTTATAGGACCTTCTGGTGGTCTCTGCTGTAC 687  
Db |||||||  
QY 179 CTGACAGCTTTGAGAAGTCTACTGCAATTTATAGGACCTTCTGGTGGTCTCTGCTGTAC 238  
Db |||||||  
QY 688 GTTTGAAGTCTGCAATCTTTGAGAACTTTGATGCAATTTATAGGACCTTCTGGTGGTCTCTGCTGTAC 747  
Db |||||||  
QY 239 GTTTGAAGTCTGCAATCTTTGAGAACTTTGATGCAATTTATAGGACCTTCTGGTGGTCTCTGCTGTAC 298  
Db |||||||  
QY 748 TTTTACAGAGAGAAACACACAGCGCAGAAATCAAGGGCCAGGCTTACTGAGGCTCTCCAGTG 807  
Db |||||||  
QY 299 TTTTACAGAGAGAAACACACAGCGCAGAAATCAAGGGCCAGGCTTACTGAGGCTCTCCAGTG 357  
Db |||||||  
QY 808 GAGGGCTCATGGTGGGACATGGAAAGAGAGGAGGCTTAGGCGCTGGGGAGCCAGTCCA 867  
Db |||||||  
QY 358 GAGGGCTCATGGTGGGACATGGAAAGAGAGGAGGCTTAGGCGCTGGGGAGCCAGTCCA 417  
Db |||||||  
QY 868 CTGAGCAGCAAGGACTGAGTGGGCTTTTGCAGGAAAGGCTTAAGAAAAAGGAAAAACC 927  
Db |||||||  
QY 418 CTGAGCAGCAAGGACTGAGTGGGCTTTTGCAGGAAAGGCTTAAGAAAAAGGAAAAACC 477  
Db |||||||  
QY 928 ATTCTTAAACACAAAGAACTGCTCAATGCTTTGGGAACTGTTTATTTGCGCTATAA 987  
Db |||||||  
QY 478 ATTCTTAAACACAAAGAACTGCTCAATGCTTTGGGAACTGTTTATTTGCGCTATAA 537  
Db |||||||  
QY 988 TGGGTCCCCAAAATGGGTAACTTAGACTTCAGAGAGAAATGAGCAGAG-AGCAAGAGGAGAA 1046  
Db |||||||  
QY 538 TGGGTCCCCAAAATGGGTAACTTAGACTTCAGAGAGAAATGAGCAGAGAGGAGAA 597  
Db |||||||  
QY 1047 ATCT-GGCTGCTCTTCAATTTTCAATCTGTTATC--TCAGGTGAGCTGGT-AGAGGGGAG 1102  
Db |||||||  
QY 598 ATCTGGGCTGCTCTTCAATTTTCAATCTGTTATCCTCAAGCTGAGCTGTTAAAGGGGAG 657  
Db |||||||  
QY 1103 ACATTAG-AAAAAATGAACAAACAACTTACTTAATGAGGTAC 1147  
Db |||||||  
QY 658 ACATTAGAAAAAATGAACAAACAACTTACTTAATGAGGTAC 703  
Db |||||||

RESULT 11

US-09-328-475C-319/c  
; Sequence 319, Application US/09328475C  
; Patent No. 6476207  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Jimmy  
; APPLICANT: Astel, Jon H.  
; APPLICANT: Carroll III, Eddie  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Ford, Donna M.  
; APPLICANT: Monahan, John E.  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Steinmann, Kathleen E.  
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT  
; FILE REFERENCE: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER  
; CURRENT APPLICATION NUMBER: US/09/328,475C  
; CURRENT FILING DATE: 1999-06-09  
; NUMBER OF SEQ ID NOS: 341  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 319  
; LENGTH: 720  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(720)  
; OTHER INFORMATION: n = A,T,C or G

US-09-328-475C-319

Query Match 30.1%; Score 582.2; DB 4; Length 720;  
Best Local Similarity 97.6%; Pred. No. 7.4e-152;  
Matches 622; Conservative 0; Mismatches 10; Indels 5; Gaps 3;  
QY 511 AAGTTGGGGCAGATTCTGAAATGGGCTAAAGACATGCAATTTTAAACCTAGGCACT 570  
Db |||||||  
QY 720 AAGTTGGGGCAGATTCTGAAAT-GGTTAAAGACATGCAATTTTAAACCTAGCAACTC 662  
Db |||||||  
QY 571 TCTTATTTCTTTCTTTTAAATAATACATAGCAATTAATCCCAATCTTATTTAAAGACCTG 630  
Db |||||||  
QY 661 ----TATTTCTTTCTTTTAAATAATACATAGCAATTAATCCCAATCTTATTTAAAGCCTG 605  
Db |||||||  
QY 631 ACAGCTTGAGAAGGCTACTGCAATTTATAGGACCTTCTGGTGGTCTTCTGCTGTACGTT 690  
Db |||||||  
QY 604 ACAGCTTGAGAAGGCTACTGCAATTTATAGGACCTTCTGGTGGTCTTCTGCTGTACGTT 545  
Db |||||||  
QY 691 TGAAGTCTGCAATCTTTGAGAACTTTTGCATGCGAGAGGTAAGAGGTATTGATTTT 750  
Db |||||||  
QY 544 TGAAGTCTGCAATCTTTGAGAACTTTTGCATGCGAGAGGTAAGAGGTATTGATTTT 485  
Db |||||||  
QY 751 CACAGAGGAAGAACACAGCGCAGAAATGAAGGCCAGGCTTACTGAGGCTGCTCCAGTGGAG 810  
Db |||||||  
QY 484 CACAGAGGAAGAACACAGCGCAGAAATGAAGGCCAGGCTTACTGAG-CTGTCTCAGTGGAG 426  
Db |||||||  
QY 811 GGCTCATGTTGGGACATGGAAAAAGAGCGACGCTTAGGCCCTGGGGAGCCAGTCCACTG 870  
Db |||||||  
QY 425 GGCTCATGTTGGGACATGGAAAAAGAGCGACGCTTAGGCCCTGGGGAGCCAGTCCACTG 366  
Db |||||||  
QY 871 AGCAAGCAAGGACTGAGTGGGCTTTTGCAGGAAAAAGGCTAAGAAAAAGGAAAAACCAT 930  
Db |||||||  
QY 365 AGCAAGCAAGGACTGAGTGGGCTTTTGCAGGAAAAAGGCTAAGAAAAAGGAAAAACCAT 306  
Db |||||||  
QY 931 CTAAACACAAACAGAACTGCTCAATGCTTTGGGAACTGTTTATTTGCTTAATGCG 990  
Db |||||||  
QY 305 CTAAACACAAACAGAACTGCTCAATGCTTTGGGAACTGTTTATTTGCTTAATGCG 246  
Db |||||||  
QY 991 GTCCCCAAAATGGGTAACTTAGACTTCAGAGAGAAATGAGCAGAGCAAGAGAGAACTCT 1050  
Db |||||||  
QY 245 GTCCCCAAAATGGGTAACTTAGACTTCAGAGAGAAATGAGCAGAGCAAGAGAGAACTCT 186  
Db |||||||  
QY 1051 GGCTGCTCTTCAATTTTCAATCTGTTATCTAGGTGAGCTGCTGAGGGGAGACATTAGA 1110  
Db |||||||  
QY 185 GGCTGCTCTTCAATTTTCAATCTGTTATCTCAGGTGAGCTGCTGAGGGGAGACATTAGA 126  
Db |||||||  
QY 1111 AAAAAATGAACAAACAAACAACTTACTTAATGAGGTAC 1147  
Db |||||||  
QY 125 AAAAAATGAACAAACAAACAACTTACTTAATGAGGTAC 89  
Db |||||||

RESULT 12

US-09-328-475C-318/c  
; Sequence 318, Application US/09328475C  
; Patent No. 6476207  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Jimmy  
; APPLICANT: Astel, Jon H.  
; APPLICANT: Carroll III, Eddie  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Ford, Donna M.  
; APPLICANT: Monahan, John E.  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Steinmann, Kathleen E.  
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT  
; FILE REFERENCE: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER  
; CURRENT APPLICATION NUMBER: US/09/328,475C  
; CURRENT FILING DATE: 1999-06-09  
; NUMBER OF SEQ ID NOS: 341  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 318  
; LENGTH: 744

; TYPE: DNA  
; ORGANISM: Homo Sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(744)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-328-475C-318

Query Match 30.0%; Score 581; DB 4; Length 744;  
Best Local Similarity 98.1%; Pred. No. 1.6e-151;  
Matches 619; Conservative 0; Mismatches 7; Indels 5; Gaps 3;

```
QY 517 GGGGGCAGATCTGAATGGGCTAAAGACATGCAATTTTAAACTAGGCACTTCTTAT 576
Db 744 GGGCCGAGATTTGAATGG--TAAAGACATGCAATTTTAAACTAG--CAACTCTTAT 689
QY 577 TTCTTCTCTTAAATAACATACATGCAATTAATCCCAATCTTATTTAAAGACCTGACAGCT 636
Db 688 TTCTTCTCTTAAATAACATACATGCAATTAATCCCAATCTTATTTAAAGACCTGACAGCT 629
QY 637 TGAGAAGGTCACTACTGCAATTTATAGGACCTTCTGGTGGTCTGCTGTTAGCTTTGAAGT 696
Db 628 TGAGAAGGTCACTACTGCAATTTATAGGACCTTCTGGTGGTCTGCTGTTAGCTTTGAAGT 569
QY 697 CTGACATCTCTGAGATCTTTCATGACAGAGAGGTAGAGTATTGGATTTTCACAGA 756
Db 568 CTGACATCTCTGAGATCTTTCATGACAGAGAGGTAGAGTATTGGATTTTCACAGA 509
QY 757 GGAAGAACACAGCGCAGAAATGAAGGCCAGGCTTACTGAGGCTGTCACAGTGGAGGCTCA 816
Db 508 GGAAGAACACAGCGCAGAAATGAAGGCCAGGCTTACTGAGGCTGTCACAGTGGAGGCTCA 450
QY 817 TGGGTGGGACATGGAAGAACAGGAGCTTAGGCCCTGGGGAGCCAGTCCACTGAGCAAG 876
Db 449 TGGGTGGGACATGGAAGAACAGGAGCTTAGGCCCTGGGGAGCCAGTCCACTGAGCAAG 390
QY 877 CAAGGACTAGTGAGGCTTTTGCAGGAAAGGCTTAGAAAGAGGAAACCATTTTAAAA 936
Db 389 CAAGGACTAGTGAGGCTTTTGCAGGAAAGGCTTAGAAAGAGGAAACCATTTTAAAA 330
QY 937 CACACAGAAACTGTCAAATGCTTTGGGAATCTGTATTATGCTTATATGGTCCCTCCC 996
Db 329 CACACAGAAACTGTCAAATGCTTTGGGAATCTGTATTATGCTTATATGGTCCCTCCC 270
QY 997 AAAATGGGTAACTAGACTTCAGAGAGAAATGAGCAGAGAGCAAGAGAAATCTGGCTGT 1056
Db 269 AAAATGGGTAACTAGACTTCAGAGAGAAATGAGCAGAGAGCAAGAGAAATCTGGCTGT 210
QY 1057 CTTTCCATTTTCATCTGTTATCTAGTCTGAGTGGTGGAGGAGACATTTAGAAAAAAA 1116
Db 209 CTTTCCATTTTCATCTGTTATCTCAGTCTGAGTGGTGGAGGAGACATTTAGAAAAAAA 150
QY 1117 TGAACACAAAAACAATTAATGAGGTAC 1147
Db 149 TGAACACAAAAACAATTAATGAGGTAC 119
```

## RESULT 13

US-09-328-475C-19

; Sequence 19, Application US/09328475C

; Patent No. 6476207

; GENERAL INFORMATION:

; APPLICANT: Zhang, Jimmy

; APPLICANT: Astel, Jon H.

; APPLICANT: Carroll III, Eddie

; APPLICANT: Endege, Wilson O.

; APPLICANT: Ford, Donna M.

; APPLICANT: Monahan, John E.

; APPLICANT: Schlegel, Robert

; APPLICANT: Steinmann, Kathleen E.

; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT

; TITLE OF INVENTION: ARE DIFFERENTIALY REGULATED IN PROSTATE CANCER

; FILE REFERENCE: 1532.002/200130.463

; CURRENT APPLICATION NUMBER: US/09/328,475C  
; CURRENT FILING DATE: 1999-06-09  
; NUMBER OF SEQ ID NOS: 341  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 19  
; LENGTH: 980  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(980)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-328-475C-19

Query Match 27.9%; Score 539.6; DB 4; Length 980;  
Best Local Similarity 92.7%; Pred. No. 5.7e-140;  
Matches 635; Conservative 0; Mismatches 42; Indels 8; Gaps 7;

```
QY 448 CTTTTTNTGTTTTTTTTTTTTTTTCCGTCCTCCCAAGCTTTATCTGCTTGGACTTTTAA 507
Db 2 CTTTTTCTTTTTTTTTTTTTTTTTTTTCCGTCCTCCCAAGCTTTATCTGCTTGGACTTTTAA 61
QY 508 AAAAAGTTTGGGGCAGATTTCTGAATGGGCTTAAAGACATGCAATTTTAAACCTAGGCA 567
Db 62 AAAAAGTTTGGGGCAGATTTCTGAAT--GGCTAAAGACATGCAATTTTAAACCTAG--C 118
QY 568 ACTTCTTATTTCTTTTCCCTTAAAAATACATAGCAATTAATCCCAATCTTATTTAAAGAC 627
Db 119 AACTCTTATTTCTTTTCCCTTAAAAATACATAGCAATTAATCCCAATCTTATTTAAAGAC 178
QY 628 CTGACAGCTTCAGAGAGTCACTACTGCAATTTATAGGACCTTCTGGTGGTCTGCTGTAC 687
Db 179 CTGACAGCTTCAGAGAGTCACTACTGCAATTTATAGGACCTTCTGGTGGTCTGCTGTAC 238
QY 688 GTTTGAAGTCTGACAAATCTTTGAGAAATCTTTCATGCAAGAGGAGGCTTAAGAGGTATGGAT 747
Db 239 GTTTGAAGTCTGACAAATCTTTGAGAAATCTTTCATGCAAGAGGAGGCTTAAGAGGTATGGAT 298
QY 748 TTTTACAGAGAGAGAACACAGCGCAGAAATGAAGGGCCAGGCTTACTGAGGCTGTCCAGTG 807
Db 299 TTTTACAGAGAGAGAACACAGCGCAGAAATGAAGGGCCAGGCTTACTGAGGCTGTCCAGTG 357
QY 808 GAGGGCTCATGGGTGGGACATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCCA 867
Db 358 GAGGGCTCATGGGTGGGACATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCCA 417
QY 868 CTGAGCAAGCAAGGAGCTGAGTG--AGCCTTTTGCAGAAAAAGGCTTAAGAAAAAGGAAAC 926
Db 418 CTGAGCAAGCAAGGAGCTGAGTGAGAGCCTTTTGCAGAAAAAGGCTTAAGAAAAAGGAAAC 477
QY 927 CATTTAAAAACAAACAAGAACTGTCAAATGCTTTGGAACTGTGTTTATGCTTATA 986
Db 478 CATTTAAAAACAAACAAGAACTGTCAAATGCTTTGGAACTGTGTTTATGCTTATA 537
QY 987 ATGGGTCCCCAAAAT--GGGTAACTTAGACTTCAGAGAGAAATGAGCAGAGAGCAAA--GGAG 1044
Db 538 ATGGGTCCCCAAAATGGGTAACCTTAGACTTCAGAGAGAAATGAGCAGAGAGCAAA--GGAG 597
QY 1045 AAATCTGGCTGCTCTCC--ATTTTCAATCTGTTATCTCAGGTGAGCTGGTAGAGGGGAGA 1103
Db 598 AAATCTGGCTGCTCTCCAAATTTTCAATCCCGTATCCAGGTGAAGCTGGGTANGAGGGG 657
QY 1104 CATTAGAAAAAATGAACACAAA 1128
Db 658 GAGANCAATNGNANAAAAATNGAA 682
```

## RESULT 14

US-09-439-313-434

; Sequence 434, Application US/09439313

; Patent No. 6329505

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

```

; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 434
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-352-616A-434

Query Match      24.2%;      Score 469.4;  DB 4;  Length 484;
Best Local Similarity 99.2%;  Pred. No. 1.2e-120;
Matches 483;  Conservative 0;  Mismatches 1;  Indels 3;  Gaps 1;

QY      1413  TTTTAAATAAGCATTTAGTGTCTCAGTCCCTACTGAGTACTCTTTCTCTCCCTCCTCTG 1472
DB      1    TTTTAAATAAGCATTTAGTGTCTCAGTCCCTACTGAGTACTCTTTCTCTCCCTCCTCTG 60

QY      1473  AATTAAATCTTTTCAACTTGC AATTGCAAGGATTAACAATTCTACTGTGATGATATTG 1532
DB      61  AATTAAATCTTTTCAACTTGC AATTGCAAGGATTAACAATTCTACTGTGATGATATTG 120

QY      1533  TGTTCGAGNGAAAGAAAGAGTGCTTTTGTGTTTAAATTTACTTTGGTTTCTGTAATCCATCT 1592
DB      121  TGTGTGCA---AAAAAAGAAAGTGCTTTGTTTAAATTTACTTTGGTTTCTGTAATCCATCT 177

QY      1593  TGTCTTTTCCCATTGGAACTAGTCAATTAACCCATCTCTGAACCTGGTAGAAAAACATCTG 1652
DB      178  TGTCTTTTCCCATTGGAACTAGTCAATTAACCCATCTCTGAACCTGGTAGAAAAACATCTG 237

QY      1653  AAGAGCTAGTCTATACAGCATCTGACAGGTGAATTGGAATGGTCTCAGAACCATTTCACCC 1712
DB      238  AAGAGCTAGTCTATCAGCATCTGACAGGTGAATTGGAATGGTCTCAGAACCATTTCACCC 297

QY      1713  AGACAGCCGTTTCTATCCTGTTTAAATAAATTTAGTTTGGTTTCTCTACATGCATAACAAA 1772
DB      298  AGACAGCCGTTTCTATCCTGTTTAAATAAATTTAGTTTGGTTTCTCTACATGCATAACAAA 357

QY      1773  CCTGCTCCAATCTGTACATAAAAGTCTGTGACTTGAAGTTTAGTCAGCACCCCCACCA 1832
DB      358  CCTGCTCCAATCTGTACATAAAAGTCTGTGACTTGAAGTTTAGTCAGCACCCCCACCA 417

QY      1833  AACTTTATTTTCTATGCTGTTTTTTTGCACATATGAGTGTGTTTGAATAAAGTACCCAT 1892
DB      418  AACTTTATTTTCTATGCTGTTTTTTTGCACATATGAGTGTGTTTGAATAAAGTACCCAT 477

QY      1893  GTCTTTA 1899
DB      478  GTCTTTA 484

Search completed: August 26, 2003, 18:32:04
Job time : 176 secs

```

Search completed: August 26, 2003, 18:32:04  
Job time : 176 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2003, 17:07:03 ; Search time 576 Seconds  
(without alignments)  
7556.328 Million cell updates/sec

Title: US-09-700-700-1  
Perfect score: 1936  
Sequence: 1 atgtgtatgccaaacttaagt.....gggcggccgcgactagtga 1936

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1517243 seqs, 1124081882 residues

Total number of hits satisfying chosen parameters: 3034486

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1836.4	94.9	2984	9	US-09-759-143-335 Sequence 335, App
2	1836.4	94.9	2984	9	US-09-780-669-335 Sequence 335, App
3	1836.4	94.9	2984	9	US-09-822-827-335 Sequence 335, App
4	1836.4	94.9	2984	10	US-09-232-880-335 Sequence 335, App
5	1836.4	94.9	2984	10	US-09-895-793-335 Sequence 335, App
6	1836.4	94.9	2984	13	US-10-012-896-335 Sequence 335, App
7	1836.4	94.9	2984	14	US-10-010-940-335 Sequence 335, App
8	1836.4	94.9	2984	14	US-10-203-823-282 Sequence 335, App
9	1645.6	85.0	3266	14	US-10-203-823-282 Sequence 335, App
10	1388.8	71.7	2051	13	US-10-202-193-334 Sequence 334, App
11	686.2	35.4	755	10	US-09-925-300-258 Sequence 258, App
12	641.2	33.1	1024	13	US-10-202-193-81 Sequence 81, Appl
13	628.4	32.5	1024	13	US-10-202-193-96 Sequence 96, Appl
14	628.2	32.4	1024	13	US-10-202-193-82 Sequence 82, Appl
15	622	32.1	1024	13	US-10-202-193-95 Sequence 95, Appl
16	615.6	31.8	1013	13	US-10-202-193-14 Sequence 14, Appl

C 17	611.8	31.6	724	10	US-09-768-827-85	Sequence 85, Appl
18	584.6	30.2	948	13	US-10-202-193-23	Sequence 23, Appl
C 19	582.2	30.1	720	13	US-10-202-193-319	Sequence 319, App
C 20	581	30.0	744	13	US-10-202-193-318	Sequence 318, App
C 21	580.6	30.0	716	10	US-09-768-827-68	Sequence 68, Appl
C 22	544.4	28.1	663	10	US-09-768-827-63	Sequence 63, Appl
C 23	539.6	27.9	980	13	US-10-202-193-19	Sequence 19, Appl
24	488	25.2	691	10	US-09-768-827-76	Sequence 76, Appl
25	477	24.6	697	10	US-09-768-827-77	Sequence 77, Appl
26	476.8	24.6	579	10	US-09-768-827-59	Sequence 59, Appl
27	469.4	24.2	484	9	US-09-759-143-434	Sequence 434, App
28	469.4	24.2	484	9	US-09-780-669-434	Sequence 434, App
29	469.4	24.2	484	9	US-09-822-827-434	Sequence 434, App
30	469.4	24.2	484	10	US-09-895-793-434	Sequence 434, App
31	469.4	24.2	484	10	US-09-895-814-434	Sequence 434, App
32	469.4	24.2	484	13	US-10-012-896-434	Sequence 434, App
33	469.4	24.2	484	14	US-10-010-940-434	Sequence 434, App
34	461.6	23.8	582	10	US-09-768-827-78	Sequence 78, Appl
35	447.4	23.1	553	10	US-09-768-827-58	Sequence 58, Appl
C 36	446.6	23.1	683	10	US-09-768-827-90	Sequence 90, Appl
C 37	426.2	22.0	731	10	US-09-768-827-55	Sequence 55, Appl
C 38	421.6	21.8	1020	13	US-10-202-193-102	Sequence 102, App
C 39	421.6	21.8	1021	13	US-10-202-193-103	Sequence 103, App
C 40	411	21.2	822	13	US-10-202-193-223	Sequence 223, App
41	378.2	19.5	497	13	US-10-202-193-222	Sequence 222, App
C 42	358.8	18.5	374	10	US-09-969-708-40	Sequence 40, Appl
C 43	350.4	18.1	366	9	US-09-759-143-115	Sequence 115, App
44	350.4	18.1	366	9	US-09-780-669-115	Sequence 115, App
45	350.4	18.1	366	9	US-09-030-606-115	Sequence 115, App

ALIGNMENTS

RESULT 1  
US-09-759-143-335  
; Sequence 335, Application US/09759143  
; Patent No. US2002022248A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John H.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C23  
; CURRENT APPLICATION NUMBER: US/09/759,143  
; CURRENT FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 934  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 335  
; LENGTH: 2984  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-759-143-335

Query Match 94.9%; Score 1836.4; DB 9; Length 2984;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1897; Conservative 0; Mismatches 12; Indels 7; Gaps 5;

QY 1 AATGGTATGCCAACTTAAGTATTTTACAGGGTGGCCCAATAGAAACAAGTGCATCTCGCTG 60  
Db |||||  
QY 1067 AATGGTATGCCAACTTAAGTATTTTACAGGGTGGCCCAATAGAAACAAGTGCATCTCGCTG 1126  
Db |||||  
QY 61 TGAATTTAAGACAAAGCTGTATATAACAGAACTCCACTGCAAGAGGNGGGCGGGCCAGGA 120  
Db |||||  
QY 1127 TGAATTTAAGACAAAGCTGTATATAACAGAACTCCACTGCAAGAGGNGGGCGGGCCAGGA 1186  
Db |||||  
QY 121 GAATCTCCGCTGTCTCAACACAGGGGCTTAAGAGGGTCTCCACTCTGCTAGGGGCT 180  
Db |||||  
QY 1187 GAATCTCCGCTGTCTCAACACAGGGGCTTAAGAGGGTCTCCACTCTGCTAGGGGCT 1246  
Db |||||  
QY 181 GTTGCAATTTTTTATTTAGTAGAAAGTGGAAAGGCTCTTCTCAACTTTTTTCCCTTGGGC 240  
Db |||||  
QY 1247 GTTGCAATTTTTTATTTAGTAGAAAGTGGAAAGGCTCTTCTCAACTTTTTTCCCTTGGGC 1306  
Db |||||  
QY 241 TGGAGAAATTTAGAAATCAGAAAGTTTCTGGAGTTTTCAGGCTATCATATATCTGTATCTT 300  
Db |||||  
QY 1307 TGGAGAAATTTAGAAATCAGAAAGTTTCTGGAGTTTTCAGGCTATCATATATCTGTATCTT 1366  
Db |||||  
QY 301 GAAAGGCAACATAATCTTCTTCCCTCTCTTTTAAATTTTGTCTCTTTCGACGAA 360  
Db |||||  
QY 1367 GAAAGGCAACATAATCTTCTTCCCTCTCTTTTAAATTTTGTCTCTTTCGACGAA 1426  
Db |||||  
QY 361 TTACTCACTAAGAGGCTTCAATTTAGTCCAGATTTTGTCTGGCTGCACCTAATCTATG 420  
Db |||||  
QY 1427 TTACTCACTAAGAGGCTTCAATTTAGTCCAGATTTTGTCTGGCTGCACCTAATCTATG 1486  
Db |||||  
QY 421 CCTCGCTTATTTAGCCCGAGATCTGGTCTTTTTTNTGNTTTTTTTTTTTCGCTCC 480  
Db |||||  
QY 1487 CCTCGCTTATTTAGCCCGAGATCTGGTCTTTTTTTT- -TTTTTTTTTTTTTTCGCTCC 1544  
Db |||||  
QY 481 CAAAGCTTATCTGTCTTGAATTTTAAAGATTTTGGGGCAGATTTCTGAATTTGGGCTA 540  
Db |||||  
QY 1545 CAAAGCTTATCTGTCTTGAATTTTAAAGATTTTGGGGCAGATTTCTGAATTTGGCTA 1603  
Db |||||  
QY 541 AAAGCATGCAATTTTAAACATAGGCACTTCTTATTTCTTTTCTTTTAAATATACATAGC 600  
Db |||||  
QY 1604 AAAGCATGCAATTTTAAACATAG- -CAACTCTTATTTCTTTTCTTTTAAATATACATAGC 1661  
Db |||||  
QY 601 ATTAATCCCAATCTCTTATTTAAAGACTGACGCTTGAAGGCTCACTACTGCAATTTAT 660  
Db |||||  
QY 1662 ATTAATCCCAATCTCTTATTTAAAGACTGACGCTTGAAGGCTCACTACTGCAATTTAT 1721  
Db |||||  
QY 661 AGGACCTTCTGGTGGTCTGTCTGTGTGAGTCTGCAATCTTGAAGTCTTGTGATCTTTGC 720  
Db |||||  
QY 1722 AGGACCTTCTGGTGGTCTGTCTGTGTGAGTCTGCAATCTTGAAGTCTTGTGATCTTTGC 1781  
Db |||||  
QY 721 ATGCAGAGGAGTAAGAGGTATTTGATTTTTCAGAGGAGAAACACAGCCGAGAAATGAAG 780  
Db |||||  
QY 1782 ATGCAGAGGAGTAAGAGGTATTTGATTTTTCAGAGGAGAAACACAGCCGAGAAATGAAG 1841  
Db |||||  
QY 781 GGCCAGGCTTACTGAGGCTGTCCAGTGGAGGGTCTATGGTGGGACATGGAAGAGGAGGC 840  
Db |||||  
QY 1842 GGCCAGGCTTACTGA-CTGTCCAGTGGAGGGCTCATGGTGGGACATGGAAGAGGAGGC 1900  
Db |||||  
QY 841 AGCCTAGGCTGGGAGGCCAGTCTCACTGAGCAAGCAAGGACTCAGTGAAGCTTTTTCG 900  
Db |||||  
QY 1901 AGCCTAGGCTGGGAGGCCAGTCTCACTGAGCAAGCAAGGACTCAGTGAAGCTTTTTCG 1960  
Db |||||  
QY 901 AGGAAAGGCTAGAAAGAGGAAACCATTTTAAACACACAGGAACTGTCCAAATGC 960  
Db |||||  
QY 1961 AGGAAAGGCTAGAAAGAGGAAACCATTTTAAACACACAGGAACTGTCCAAATGC 2020  
Db |||||  
QY 961 TTTGGGAATCTGTGTTTATTTGCTTATATGTTGGTCCCCAAATGGTAACTAGACTTCAGA 1020  
Db |||||  
QY 2021 TTTGGGAATCTGTGTTTATTTGCTTATATGTTGGTCCCCAAATGGTAACTAGACTTCAGA 2080  
Db |||||  
QY 1021 GAGAAATGAGAGAGCAAGAGGAAATCTGGCTGTCTTCAATTTTCACTTCTGTATCT 1080  
Db |||||  
QY 2081 GAGAAATGAGAGAGCAAGAGGAAATCTGGCTGTCTTCAATTTTCACTTCTGTATCT 2140  
Db |||||  
QY 1081 CAGGTGAGCTGTGTAGAGGGGAGACATTAGAAAAAATAGAAACAACAAACAAATTTACTAAT 1140  
Db |||||

Db 2141 CAGGTGAGCTGTGTAGAGGGGAGACATTAGAAAAAATAGAAACAACAAATTTACTAAT 2200  
QY |||||  
Db 1141 GAGGTACGCTAGGCGCTGGAGTCTCTTGACTCCACTACTTAATTCGGTTTGTAGTGA 1200  
QY |||||  
Db 2201 GAGGTACGCTAGGCGCTGGAGTCTCTTGACTCCACTACTTAATTCGGTTTGTAGTGA 2260  
QY |||||  
QY 1201 CCTTCAATTTTCTTTTATTTAGAGGGCCAGCTTACTGTGTGGTGGCAAAATTTGCCAAT 1260  
Db |||||  
QY 2261 CCTTCAATTTTCTTTTATTTAGAGGGCCAGCTTACTGTGTGGTGGCAAAATTTGCCAAT 2320  
QY |||||  
QY 1261 AGTTAATAGAAAGTTGGCCAAATTTTCAAGGCTTCTGTGTGGTGGTCCCAATTTGCA 1320  
Db |||||  
QY 2321 AGTTAATAGAAAGTTGGCCAAATTTTCAAGGCTTCTGTGTGGTGGTCCCAATTTGCA 2380  
QY |||||  
QY 1321 ATGTTCAATGCGGCTGCTGCTGACACCGGAGTACTAGCCAGCAGCAAAAGCAGG 1380  
Db |||||  
QY 2381 ATGTTCAATGCGGCTGCTGCTGACACCGGAGTACTAGCCAGCAGCAAAAGCAGG 2440  
QY |||||  
QY 1381 TAGCCTGAATTTGCTTCTGCTCTTTTACATTTCTTTTAAATTAAGCATTTAGTGTCTCAGTC 1440  
Db |||||  
QY 2441 TAGCCTGAATTTGCTTCTGCTCTTTTACATTTCTTTTAAATTAAGCATTTAGTGTCTCAGTC 2500  
QY |||||  
QY 1441 CCTACTGAGTACTCTTCTCTCCCTCTCTCTGAAATTTTCTTTCAACTTGCATTTTGC 1500  
Db |||||  
QY 2501 CTTACTGAGTACTCTTCTCTCCCTCTCTGAAATTTTCTTTCAACTTGCATTTTGC 2560  
QY |||||  
QY 1501 AAGGATTAACATTTTCTGCTGATGATATTTGTTGAGGAGGAGGAGGAGGAGGAGG 1560  
Db |||||  
QY 2561 AAGGATTAACATTTTCTGCTGATGATATTTGTTGAGGAGGAGGAGGAGGAGGAGG 2619  
QY |||||  
QY 1561 TGTTTAAATTAATTTGTTTGTGAAATCTTGTGCTTTTCCCATTTGAACTAGTCAAT 1620  
Db |||||  
QY 2620 TGTTTAAATTAATTTGTTTGTGAAATCTTGTGCTTTTCCCATTTGAACTAGTCAAT 2679  
QY |||||  
QY 1621 AACCATCTCTGAACTGTAGAAAAAATCTGAGAGCTAGTCTATAGCATCTGACAGG 1680  
Db |||||  
QY 2680 AACCATCTCTGAACTGTAGAAAAAATCTGAGAGCTAGTCTATAGCATCTGACAGG 2739  
QY |||||  
QY 1681 TGAATTTGATGTTTCTAGAACCAATTTCCACAGACGCTGTTTCTATCTCTGTTAATA 1740  
Db |||||  
QY 2740 TGAATTTGATGTTTCTAGAACCAATTTCCACAGACGCTGTTTCTATCTCTGTTAATA 2799  
QY |||||  
QY 1741 AATTAGTTTGGTCTCTACATGATGATCAACCACTGCTCAATCTGTCAATAAAGTC 1800  
Db |||||  
QY 2800 AATTAGTTTGGTCTCTACATGATGATCAACCACTGCTCAATCTGTCAATAAAGTC 2859  
QY |||||  
QY 1801 TGTGACTTGAAGTTTGTAGTCAACCCCAACCACTTTTATTTTCTATGTTTGTGCA 1860  
Db |||||  
QY 2860 TGTGACTTGAAGTTTGTAGTCAACCCCAACCACTTTTATTTTCTATGTTTGTGCA 2919  
QY |||||  
QY 1861 ACATATGAGTCTTTTGAATTAAGTACCCTGTTTATTAATAAATAAATAAATAA 1916  
Db |||||  
QY 2920 ACATATGAGTCTTTTGAATTAAGTACCCTGTTTATTAATAAATAAATAAATAA 2975  
Db |||||

## RESULT 2

US-09-780-669-335  
; Sequence 335, Application US/09780669  
; Patent No. US2002005197A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Devin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.



```

Db      2860  TGTGACTTGAAGTTTGTAGTCAGACCCACCCACCAACTTTATTTTCTATGTGTTTTTGCA 2919
Qy      1861  ACATATGAGTGTGTTTGAATAAAGTACCCATGTCTTTATTAATAAATAAATAAATAA 1916
Db      2920  ACATATGAGTGTGTTTGAATAAAGTACCCATGTCTTTATTAATAAATAAATAAATAA 2975

RESULT 3
US-09-822-827-335
; Sequence 335, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 335
; LENGTH: 2984
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-822-827-335

Query Match      94.9%; Score 1836.4; DB 9; Length 2984;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 189; Conservative 0; Mismatches 12; Indels 7; Gaps 5;

Qy      1  AATGGTATGCCAACTTAAAGTATTATACAGGGTGGCCCAATAGACACAGATGCATCGCTG 60
Db      1067  AATGGTATGCCAACTTAAAGTATTATACAGGGTGGCCCAATAGACACAGATGCATCGCTG 1126
Qy      61  TGATTTTAAAGACAGCTGTATATAAAGAGTCTTCTCACTGCAAGAGGCGGCGGCGCAGGA 120
Db      1127  TGATTTTAAAGACAGCTGTATATAAAGAGTCTTCTCACTGCAAGAGGCGGCGGCGCAGGA 1186
Qy      121  GAATCTCCGCTTGTCCAAAGACAGGGGCTTAAGAGGGTCTCCACATGCTGTAGGGGCT 180
Db      1187  GAATCTCCGCTTGTCCAAAGACAGGGGCTTAAGAGGGTCTCCACATGCTGTAGGGGCT 1246
Qy      181  GTTGCAATTTTATTAGTAGAAGTGAAGAGGCTCTTCTCACTTTTTCCTTGGG 240
Db      1247  GTTGCAATTTTATTAGTAGAAGTGAAGAGGCTCTTCTCACTTTTTCCTTGGG 1306
Qy      241  TGGAGAAATTTAGAAATCAGAACTTTCCGAGGTTTTCAGGCTATCATATATCTGATCT 300
Db      1307  TGGAGAAATTTAGAAATCAGAACTTTCCGAGGTTTTCAGGCTATCATATATCTGATCT 1366
Qy      301  GAAAGGCAACATAATCTCTCCCTCTTTTAAATTTTGTCTTTTTCCTTTGCGAA 360
Db      1367  GAAAGGCAACATAATCTCTCCCTCTTTTAAATTTTGTCTTTTTCCTTTGCGAA 1426
Qy      361  TTACTCACTAAGGGCTTCATTTTATGTCAGATTTTATGTCGGCTGCACTTAATG 420
Db      1427  TTACTCACTAAGGGCTTCATTTTATGTCAGATTTTATGTCGGCTGCACTTAATG 1486
Qy      421  CCTCGCTATTATAGCCGAGATCTGCTCTTTTNTGNTTTTTTTTTTTTTTTTTTTT 480
Db      1487  CCTCGCTATTATAGCCGAGATCTGCTCTTTTNTGNTTTTTTTTTTTTTTTTTTTT 1544
Qy      481  CAAAGCTTTATCTGCTTCACTTTTAAAGAGTTTGGGCGAGATCTGAAATGGGCTA 540
Db      1545  CAAAGCTTTATCTGCTTCACTTTTAAAGAGTTTGGGCGAGATCTGAAATGGGCTA 1603
Qy      541  AAAGACATCATTTTAAACTAGGCACTTTTATTTCTTTCTTTTAAATAATACATAGC 600
Db      1604  AAAGACATCATTTTAAACTAG--CAACTCTTATTTCTTTCTTTTAAATAATACATAGC 1661
Qy      601  ATTAATCCAAATCCTATTAAAGACCTGACAGCTGAGAGTCACTGCTATTTAT 660

```

```

Db      1662  ATTAATCCAAATCCTATTATTAAGACCTGACAGCTTGAGAGGTCACCTACTGCAATTTAT 1721
Qy      661  AGGACCTTCTGGTGGTCTCTGCTAGCTTTGAAGTCTGCAATCTTCTGAGAACTTTTGC 720
Db      1722  AGGACCTTCTGGTGGTCTCTGCTAGCTTTGAAGTCTGCAATCTTCTGAGAACTTTTGC 1781
Qy      721  ATGCAGAGGAGGTAAGAGGTAATTTTTCAGAGGAGAAACACAGCGCAGAAATGAAG 780
Db      1782  ATGCAGAGGAGGTAAGAGGTAATTTTTCAGAGGAGAAACACAGCGCAGAAATGAAG 1841
Qy      781  GGCACAGGCTTACTGAGGCTGTCAGTGGAGGCTCATGGGTGGGACATGGAAAAAGAGC 840
Db      1842  GGCACAGGCTTACTGA--GCTGTCCAGTGGAGGCTCATGGGTGGGACATGGAAAAAGAGC 1900
Qy      841  AGCTTAGGCTTGGGGAGCCCACTGCACTGAGCAAGAGGAGCTGAGTGGAGCTTTTGC 900
Db      1901  AGCTTAGGCTTGGGGAGCCCACTGCACTGAGCAAGAGGAGCTGAGTGGAGCTTTTGC 960
Qy      901  AGGAAAAGGCTAAGAAAAGGAAAACCAATTTTAAACAACAAGAACTGTCCAAATGC 960
Db      1961  AGGAAAAGGCTAAGAAAAGGAAAACCAATTTTAAACAACAAGAACTGTCCAAATGC 2020
Qy      961  TTTGGGAATGTTTATTTGCTTATAATGGTCCCAAAATGGGTAACCTAGACTTCAGA 1020
Db      2021  TTTGGGAATGTTTATTTGCTTATAATGGTCCCAAAATGGGTAACCTAGACTTCAGA 2080
Qy      1021  GAGAAATGAGCAGAGCAAGGAGAAATCTGGCTGCTCTTCCATTTTCACTTCTGTTATCT 1080
Db      2081  GAGAAATGAGCAGAGCAAGGAGAAATCTGGCTGCTCTTCCATTTTCACTTCTGTTATCT 2140
Qy      1081  CAGGTAGCTGTTAGAGGGGAGACATTTAGAAAAAAATGAAACAACAACAATTAATAAT 1140
Db      2141  CAGGTAGCTGTTAGAGGGGAGACATTTAGAAAAAAATGAAACAACAACAATTAATAAT 2200
Qy      1141  GAGGTAGCTGTTAGGGCTTGGAGTCTCTTGAATCCACTTAAATTCGTTAGTGAGAAA 1200
Db      2201  GAGGTAGCTGTTAGGGCTTGGAGTCTCTTGAATCCACTTAAATTCGTTAGTGAGAAA 2260
Qy      1201  CCTTTCAATTTCTTTTATTTAGAGGGGCGAGCTTACTGTTGGTGCAAAATGGCAACAT 1260
Db      2261  CCTTTCAATTTCTTTTATTTAGAGGGGCGAGCTTACTGTTGGTGCAAAATGGCAACAT 2320
Qy      1261  AAGTTAATAGAAAGTTGGCCCAATTTTCAACCCATTTTCTGGTGGCTCCCACTGCA 1320
Db      2321  AAGTTAATAGAAAGTTGGCCCAATTTTCAACCCATTTTCTGGTGGCTCCCACTGCA 2380
Qy      1321  ATGTTCAATGCCACGCTGCTGACACCGGAGTACTAGCCAGCAGCAAAAGGCGAGG 1380
Db      2381  ATGTTCAATGCCACGCTGCTGACACCGGAGTACTAGCCAGCAGCAAAAGGCGAGG 2440
Qy      1381  TAGCCTGAATTTCTTCTGCTTTTACATTTTCTTTTAAATAAGCAATTTAGTGTCAAGTC 1440
Db      2441  TAGCCTGAATTTCTTCTGCTTTTACATTTTCTTTTAAATAAGCAATTTAGTGTCAAGTC 2500
Qy      1441  CCTACTGAGTACTCTTCTCTCCCTCTCTCTGAATTTTAAATTTCTTCACTGCAATTTGC 1500
Db      2501  CCTACTGAGTACTCTTCTCTCCCTCTCTCTGAATTTTAAATTTCTTCACTGCAATTTGC 2560
Qy      1501  AAGGATTACACATTTTCACTGTGATATATGTTGCGAGGAAAGAAAAAGTGTCTT 1560
Db      2561  AAGGATTACACATTTTCACTGTGATATATGTTGTTGCA--AAAAAAGAAAAAGTGTCTT 2619
Qy      1561  TGTTTAAATTTACTGTTGTTGTAATCCATTTTGTCTTTTCCCACTTGGAACTAGTCATT 1620
Db      2620  TGTTTAAATTTACTGTTGTTGTAATCCATTTTGTCTTTTCCCACTTGGAACTAGTCATT 2679
Qy      1621  AACCCATCTCTGAACCTGTTAGAAACATCTGAGAGTGTCTATCAGCATCTGCACAGG 1680
Db      2680  AACCCATCTCTGAACCTGTTAGAAACATCTGAGAGTGTCTATCAGCATCTGCACAGG 2739
Qy      1681  TGAATTTGATGTTCTCAGAACCAATTTCCAGCAGCAGCTGTTCTATCTCTGTTTATA 1740
Db      2740  TGAATTTGATGTTCTCAGAACCAATTTCCAGCAGCAGCTGTTCTATCTCTGTTTATA 2799

```

Qy	1741	AATTAGTTTGGGTTCTCTACATGCATAACAAACCCTGCTCAATCTCTGTCAATAAAAGTC	1800
Db	2800	AATTAGTTTGGGTTCTCTACATGCATAACAAACCCTGCTCAATCTCTGTCAATAAAAGTC	2859
Qy	1801	TGTGACTTTGAAGTTTGTAGTCAGCACCCCCACCACCAACTTTATTTTTCTATGTGTTTTTGTCA	1860
Db	2860	TGTGACTTTGAAGTTTGTAGTCAGCACCCCCACCACCAACTTTATTTTTCTATGTGTTTTTGTCA	2919
Qy	1861	ACATATGAGTGTTTTTGAAAAATAAAGTACCCATGCTTTTATTTAAAAANAAAAA	1916
Db	2920	ACATATGAGTGTTTTTGAAAAATAAAGTACCCATGCTTTTATTTAGAAAAA	2975

```

RESULT 4
US-09-232-880-335
; Sequence 335, Application US/09232880
; Publication No. US20020182596A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF
; TITLE OF INVENTION: PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.428C6
; CURRENT APPLICATION NUMBER: US/09/232,880
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ.ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ.ID NO 335
; LENGTH: 2984
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-232-880-335

```

Query Match	94.98;	Score 1836.4;	DB 10;	Length 2984;
Best Local Similarity	99.0%;	Pred. No. 0;		
Matches 1897;	Conservative 0;	Mismatches 12;	Indels 7;	Gaps 5;
QY 1	AATGGTATGCCAACTTAAGTATTTACAGGGTGGGCCAAATAGAACAGATGCATCGCTG	60		
DB 1067	AATGGTATGCCAACTTAAGTATTTACAGGGTGGGCCAAATAGAACAGATGCATCGCTG	1126		
QY 61	TGATTTTAAGACAAGCTGTATTAACAGAACTCCACTGCAAGAGGNGGGCCGGGCCAGGA	120		
DB 1127	TGATTTTAAGACAAGCTGTATTAACAGAACTCCACTGCAAGAGGNGGGCCGGGCCAGGA	1186		
QY 121	GAATCTCCGCTGTGTCGAAGACAGGGGCTTAAGGAGGGTCTCCACACTGCTGTAGGGGCT	180		
DB 1187	GAATCTCCGCTGTGTCGAAGACAGGGGCTTAAGGAGGGTCTCCACACTGCTGTAGGGGCT	1246		
QY 181	GTTGCAATTTTTTATTAGTAGAAGTGGAAAGGCCTCTCTCAACTTTTTTCCCTTGGGC	240		
DB 1247	GTTGCAATTTTTTATTAGTAGAAGTGGAAAGGCCTCTCTCAACTTTTTTCCCTTGGGC	1306		
QY 241	TGAGAGATTTTAGAATCAGAGATTTCTGGAGTTTTAGGCTATCATATACTGTATCCT	300		
DB 1307	TGAGAGATTTTAGAATCAGAGATTTCTGGAGTTTTAGGCTATCATATACTGTATCCT	1366		
QY 301	GAAGGGCAACATAATCTTCCCTCCCTCTTTTAAAAATTTTGTTGTTCTTTTTGCAGCAA	360		
DB 1367	GAAGGGCAACATAATCTTCCCTCCCTCTTTTAAAAATTTTGTTGTTCTTTTTGCAGCAA	1426		
QY 361	TTACTCACTAAAGGGCTTCATTTTAGTCCAGATTTTATAGTCTGGCTGCACCTAACTTATG	420		
DB 1427	TTACTCACTAAAGGGCTTCATTTTAGTCCAGATTTTATAGTCTGGCTGCACCTAACTTATG	1486		
QY 421	CCTCGCTTATTTAGCCCGAGATCTGGTCTTTTNTGTTNTTTTTTTTTTTTTCGGTCTCCC	480		
DB 1487	CCTCGCTTATTTAGCCCGAGATCTGGTCTTTTNTGTTNTTTTTTTTTTTTTCGGTCTCCC	1544		
QY 481	CAAAAGCTTTATCTGTCTTGACTTTTTTAAAAAAGTTTGGGGGCAGATTTCTGAAATTGGGCTA	540		

1545	DB	CAAAGCTTTATCTGTCTGTGACCTTTTAAAAAAGTTTGGGGGCGAGATTCCTGAATT-GGCTA	1601
541	QY	AAAGACATGCATTTTTTAAAACTAGGCAACTTCCTATTCTTTCTCTTTTAAAAATACATAGC	600
1604	DB	AAAGACATGCATTTTTTAAAACTAG--CAACTCTTATTCTTTCTCTTTTAAAAATACATAGC	1661
601	QY	ATTAAATCCCAATCCTATTATAAGACCTGCACAGCTTGAGAAGGTCACTACTGCATTTTAT	660
1662	DB	ATTAAATCCCAATCCTATTATAAGACCTGCACAGCTTGAGAAGGTCACTACTGCATTTTAT	1721
661	QY	AGGACCTTCTGGTGGTTCTGCTGTACGTTTGAAGTCTGACAATCCTTGAGAACTTTTGC	720
1722	DB	AGGACCTTCTGGTGGTTCTGCTGTACGTTTGAAGTCTGACAATCCTTGAGAACTTTTGC	1781
721	QY	ATCGAGAGGAGGTAAGAGGTAATTGGATTTTTTCACAGAGGAGAAACACAGCGCAGAAATGAAG	780
1782	DB	ATCGAGAGGAGGTAAGAGGTAATTGGATTTTTTCACAGAGGAGAAACACAGCGCAGAAATGAAG	1841
781	QY	GGCCAGGCTTACTGAGGCTGTCCAGTGGAGGGCTCATGGGTGGGAATGGAAGAAGGC	840
1842	DB	GGCCAGGCTTACTGGA-GCTGTCCAGTGGAGGGCTCATGGGTGGGAATGGAAGAAGGC	1900
841	QY	AGCCTAGGCCCTGGGAGGCCAGTCCACTGAGCAAGCAAGGGACTGAGTGAGCCTTTTGC	900
1901	DB	AGCCTAGGCCCTGGGAGGCCAGTCCACTGAGCAAGCAAGGGACTGAGTGAGCCTTTTGC	1960
901	QY	AGGAAAAGGCTAAGAAAAAGGAAAAACATTTCTAAAAACAACAAGAAATGTCCTCAAAATGC	960
1961	DB	AGGAAAAGGCTAAGAAAAAGGAAAAACATTTCTAAAAACAACAAGAAATGTCCTCAAAATGC	2020
961	QY	TTTGGGAATCTGTCTTTATTGCTTAAATGGGTCCCCAAAATGGGTAACTCTAGACTTTCAGA	1020
2021	DB	TTTGGGAATCTGTCTTTATTGCTTAAATGGGTCCCCAAAATGGGTAACTCTAGACTTTCAGA	2080
1021	QY	GAGAAATGACGACAGACAAAGGAAATCTGGCTGTCTTCCATTTTTCATCTCTGTTATCT	1080
2081	DB	GAGAAATGACGACAGACAAAGGAAATCTGGCTGTCTTCCATTTTTCATCTCTGTTATCT	2140
1081	QY	CAGGTGAGCTGGTAGAGGGGAGACATTAGAAAAAATGAAAAACAACAACAATTACTAAT	1140
2141	DB	CAGGTGAGCTGGTAGAGGGGAGACATTAGAAAAAATGAAAAACAACAACAATTACTAAT	2200
1141	QY	GAGGTACGTGAGGCTCGGAGTCTCTTGACTCCACTACTTAATTCGGTTTAGTGAGAAA	1200
2201	DB	GAGGTACGTGAGGCTCGGAGTCTCTTGACTCCACTACTTAATTCGGTTTAGTGAGAAA	2260
1201	QY	CCTTTCAATTTCTTTTATTAGAGGGCCAGCTTACTGTGGTGGCAAAATTTGCCAACAT	1260
2261	DB	CCTTTCAATTTCTTTTATTAGAGGGCCAGCTTACTGTGGTGGCAAAATTTGCCAACAT	2320
1261	QY	AAAGTAAATAGAAATTTGGCCAAATTTTCAACCCCAATTTCTGTGGTTGGGCTCCACATGCA	1320
2321	DB	AAAGTAAATAGAAATTTGGCCAAATTTTCAACCCCAATTTCTGTGGTTGGGCTCCACATGCA	2380
1321	QY	ATGTTTCAATGCCACGTGCTGCTGACACCGACGGAGTACTAGCCAGCACAAAAAGGCAGGG	1380
2381	DB	ATGTTTCAATGCCACGTGCTGCTGACACCGACGGAGTACTAGCCAGCACAAAAAGGCAGGG	2440
1381	QY	TAGCCTGAATGCTTTCTGCTTTTACATTTCTTTTAAAAATAGAAATTTAGTGCTCAGTC	1440
2441	DB	TAGCCTGAATGCTTTCTGCTTTTACATTTCTTTTAAAAATAGAAATTTAGTGCTCAGTC	2500
1441	QY	CTTACTGAGTACTCTTTTCTCTCCCTCCCTGAAATTTAAATCTTTTCAACTGCAATTTGC	1500
2501	DB	CTTACTGAGTACTCTTTTCTCTCCCTCCCTGAAATTTAAATCTTTTCAACTGCAATTTGC	2560
1501	QY	AAGGATTAACATTTTCACTGTGATGATATATGTTTGCAGNAGAAAAAGTGTCTTT	1560
2561	DB	AAGGATTAACATTTTCACTGTGATGATATATGTTTGCAGNAGAAAAAGTGTCTTT	2619
1561	QY	TGTTTAAAAATTACTGGTTTGAATCCATCTGCTTTTTCCTCCCATTTGGAACTAGTCAAT	1620
2620	DB	TGTTTAAAAATTACTGGTTTGAATCCATCTGCTTTTTCCTCCCATTTGGAACTAGTCAAT	2679

```

QY 1621 AACCATCTCTGAAGTGTAGAAAACATCTGAAGAGCTAGTCTATCAGCATCTCAGAGG 1680
Db 2680 AACCACTCTGAAGTGTAGAAAACATCTGAAGAGCTAGTCTATCAGCATCTCAGAGG 2739
QY 1681 TGAATTGATGTTCTCAGAACCAATTTACCCAGAGCAGCTGTTCTATCTCTGTTTAATA 1740
Db 2740 TGAATTGATGTTCTCAGAACCAATTTACCCAGAGCAGCTGTTCTATCTCTGTTTAATA 2799
QY 1741 AATTAGTTTGGGTTCTCTACATGATCAATAACAAACCTGCTCCAACTGTGCATATAAAGTC 1800
Db 2800 AATTAGTTTGGGTTCTCTACATGATCAATAACAAACCTGCTCCAACTGTGCATATAAAGTC 2859
QY 1801 TGTGACTTGAAGTTAGTCAGCACCCCAACCAAACTTTATTTTCTATGTTGTTTGA 1860
Db 2860 TGTGACTTGAAGTTAGTCAGCACCCCAACCAAACTTTATTTTCTATGTTGTTTGA 2919
QY 1861 ACATATGAGTGTGTTTGAATAAAGTACCAATGCTTTTATTAATAAATAAATAA 1916
Db 2920 ACATATGAGTGTGTTTGAATAAAGTACCAATGCTTTTATTAATAAATAAATAA 2975

```

## RESULT 5

```

US-09-895-793-335
; Sequence 335, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitchem, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 335
; LENGTH: 2984
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-895-793-335

```

```

Query Match 94.9%; Score 1836.4; DB 10; Length 2984;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1897; Conservative 0; Mismatches 12; Indels 7; Gaps 5;
QY 1 AATGGTATGCCAACTTAACTATTACAGGGTGGCCCAATAGAACAGATCGACTCGCTG 60
Db 1067 AATGGTATGCCAACTTAACTATTACAGGGTGGCCCAATAGAACAGATCGACTCGCTG 1126
QY 61 TGAATTTAGACAAGCTGTATTAACAGAACTCCACTGCAAGAGGCGGCGGCCAGGA 120
Db 1127 TGAATTTAGACAAGCTGTATTAACAGAACTCCACTGCAAGAGGCGGCGGCCAGGA 1186

```

```

QY 121 GAATCTCCGCTTGTCCRAGACAGGGGCTAAGAGGGTCTCCACACTGCTGTAGGGCT 180
Db 1187 GAATCTCCGCTTGTCCRAGACAGGGGCTAAGAGGGTCTCCACACTGCTGTAGGGCT 1246
QY 181 GTTGCAATTTTATTAAGTAGAAGTGAAAGGCTCTTCTCAACTTTTCCCTTGGGC 240
Db 1247 GTTGCAATTTTATTAAGTAGAAGTGAAAGGCTCTTCTCAACTTTTCCCTTGGGC 1306
QY 241 TGGAGAAATTTAGAAATCAGAGTTTCTGAGATTTTCAGGCTATCATATATCTGTATCT 300
Db 1307 TGGAGAAATTTAGAAATCAGAGTTTCTGAGATTTTCAGGCTATCATATATCTGTATCT 1366
QY 301 GAAAGGCAACATAATTTCTTCTTCCCTCTTTTAAATTTTGTGTCTCTTTTGGCAGAA 360
Db 1367 GAAAGGCAACATAATTTCTTCTTCCCTCTTTTAAATTTTGTGTCTCTTTTGGCAGAA 1426
QY 361 TTACTCACTAAAGGGCTTCATTTTATAGTCAGATTTTATAGTCGCTGACCTAACTTATG 420
Db 1427 TTACTCACTAAAGGGCTTCATTTTATAGTCAGATTTTATAGTCGCTGACCTAACTTATG 1486
QY 421 CCTCGCTTATTTAGCCCGAGATCTGGTCTTTTNTGNTTNTTNTTNTTNTTNTTNTTNTT 480
Db 1487 CCTCGCTTATTTAGCCCGAGATCTGGTCTTTTNTTNTTNTTNTTNTTNTTNTTNTTNT 1544
QY 481 CAAAGCTTTATCTGTCTTGTACTTTTAAAGTTTGGGGGCGAGATTTCTGAATTTGGGCTA 540
Db 1545 CAAAGCTTTATCTGTCTTGTACTTTTAAAGTTTGGGGGCGAGATTTCTGAATTTGGGCTA 1603
QY 541 AAAGACATGCAATTTTAAACTAGGCAACTCTTATTTCTTCTTCTTCTTCTTCTTCTTCTT 600
Db 1604 AAAGACATGCAATTTTAAACTAGGCAACTCTTATTTCTTCTTCTTCTTCTTCTTCTTCT 1661
QY 601 ATTAATCCCAATCTTATTTAAAGCTGACAGCTTGAGAGGTGACCTACTGCAATTTAT 660
Db 1662 ATTAATCCCAATCTTATTTAAAGCTGACAGCTTGAGAGGTGACCTACTGCAATTTAT 1721
QY 661 AGGACCTTCTGGTGTCTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGT 720
Db 1722 AGGACCTTCTGGTGTCTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGT 1781
QY 721 ATGACAGAGGAGTAAGAGGTATTGGAATTTTACAGAGGAAGAACACAGCGCAGAAATGAAG 780
Db 1782 ATGACAGAGGAGTAAGAGGTATTGGAATTTTACAGAGGAAGAACACAGCGCAGAAATGAAG 1841
QY 781 GGCACGCTTACTGAGGCTGTCCAGTGGAGGCTCATCGGTGGGAGGCTCATGGTGGGAGCAT 840
Db 1842 GGCACGCTTACTGGA-GCTGTCCAGTGGAGGCTCATGGTGGGAGCATGGAAGAGGAGG 1900
QY 841 AGCTAGGCGCTTGGGAGGCGCCAGTCCACTGAGGAGCAAGCAAGGAGTGTGAGCTTTTTC 900
Db 1901 AGCTAGGCGCTTGGGAGGCGCCAGTCCACTGAGGAGCAAGCAAGGAGTGTGAGCTTTTTC 1960
QY 901 AGGAAAGGCTAAGAAAAGGAAACCAATTTCTAAACACAAAGAACTGTCCAATGC 960
Db 1961 AGGAAAGGCTAAGAAAAGGAAACCAATTTCTAAACACAAAGAACTGTCCAATGC 2020
QY 961 TTTGGAACTGTGTTTATTTGCTTATATGGTCCCAAAATGGGTAACTAGACTTTCAGA 1020
Db 2021 TTTGGAACTGTGTTTATTTGCTTATATGGTCCCAAAATGGGTAACTAGACTTTCAGA 2080
QY 1021 CAGATGACAGAGCAAGAGGAGAAATCTGGTGTCTTCTTCCATTTTCTTCTTCTTCTTCT 1080
Db 2081 GAGATGACAGAGCAAGAGGAGAAATCTGGTGTCTTCTTCCATTTTCTTCTTCTTCTTCT 2140
QY 1081 CAGTGTAGCTGTGTAGAGGGGAGACATTAGAAAATAAGAAACAAACAACTACTAAT 1140
Db 2141 CAGTGTAGCTGTGTAGAGGGGAGACATTAGAAAATAAGAAACAAACAACTACTAAT 2200
QY 1141 GAGTGTAGCTGTGTAGGCTGTGGAGTCTCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200
Db 2201 GAGTGTAGCTGTGTAGGCTGTGGAGTCTCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 2260

```

1201 CCTTCAATTTCTTTTATTAGAGGCGCAGCTTACTGTGTGGGCAAAATGCCAACAT 1260  
Db CCTTCAATTTCTTTTATTAGAGGCGCAGCTTACTGTGTGGGCAAAATGCCAACAT 2320  
QY AGGTTAATAGAAAGTTGGCAATTTCCACCCCATTTCTGTGGTTTGGGCTCCACATTGCA 1320  
Db AGGTTAATAGAAAGTTGGCAATTTCCACCCCATTTCTGTGGTTTGGGCTCCACATTGCA 2380  
QY ATGTTCAATGCCAGTGCTGCTGACACCGACCGAGTACTAGCCAGCACAAGGCGAGG 1380  
Db ATGTTCAATGCCAGTGCTGCTGACACCGACCGAGTACTAGCCAGCACAAGGCGAGG 2440  
QY TAGCCTGAATGTGTTCTGCTCTTTTACATTTCTTTTAAATAAGCAATTTAGTGTCTCAGTC 1440  
Db TAGCCTGAATGTGTTCTGCTCTTTTACATTTCTTTTAAATAAGCAATTTAGTGTCTCAGTC 2500  
QY CCTACTGAGTACTCTTTCTCTCCCTCCCTCTGAAATTTAAATCTTTTCAACTTGCATTTGC 1500  
Db CCTACTGAGTACTCTTTCTCTCCCTCCCTCTGAAATTTAAATCTTTTCAACTTGCATTTGC 2560  
QY AAGGATTACACATTTCACTGTGATATATTTGTGTGCAAGGAAAGAAAGTGTCTTT 1560  
Db AAGGATTACACATTTCACTGTGATATATTTGTGTGCA-AAAAAATAAGTGTCTTT 2619  
QY TGTTTAAATTTACTGGTTTGAATCCATCTTTTCTTTTCCCATTTGAACTAGTCAAT 1620  
Db TGTTTAAATTTACTGGTTTGAATCCATCTTTTCTTTTCCCATTTGAACTAGTCAAT 2679  
QY ACCCATCTCTGAACCTGTAGAAAACATCTGAAGAGCTAGTCTATCAGCATCTGCACAG 1680  
Db ACCCATCTCTGAACCTGTAGAAAACATCTGAAGAGCTAGTCTATCAGCATCTGCACAG 2739  
QY TGAATTTGATGTTCTCAGAACCATTTACCCAGACAGCTGTTTCTATCTCTGTTTAAATA 1740  
Db TGAATTTGATGTTCTCAGAACCATTTACCCAGACAGCTGTTTCTATCTCTGTTTAAATA 2799  
QY AATTAGTTGGGTTCTCTACATGCAATAACAAACCTGCTCCAACTCTGTACATAAAAGTC 1800  
Db AATTAGTTGGGTTCTCTACATGCAATAACAAACCTGCTCCAACTCTGTACATAAAAGTC 2859  
QY TGTGACTCTGAAGTTTGTAGTACAGACCCCAACCACTTTATTTCTATGTTTGTGCA 1860  
Db TGTGACTCTGAAGTTTGTAGTACAGACCCCAACCACTTTATTTCTATGTTTGTGCA 2919  
QY ACATATGAGTGTGTTGAAATAAAGTACCCATGCTTTTATTTAAATAAATAAATAA 1916  
Db ACATATGAGTGTGTTGAAATAAAGTACCCATGCTTTTATTTAGAAAAAATAAATAA 2975

RESULT 6  
US-09-895-814-335  
; Sequence 335, Application US/09895914  
; Publication No. US20020193296A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darriack  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C26  
; CURRENT APPLICATION NUMBER: US/09/895,814  
; NUMBER OF SEQ ID NOS: 990  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 335  
; LENGTH: 2984  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-895-814-335  
Query Match 94.9%; Score 1836.4; DB 10; Length 2984;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1897; Conservative 0; Mismatches 12; Indels 7; Gaps 5;  
QY 1 AATGGTATGCCAACTTAAGTATTATTACAGGGTGGCCCAATAGAACAAAGATGCATCGCTG 60  
Db 1067 AATGGTATGCCAACTTAAGTATTATTACAGGGTGGCCCAATAGAACAAAGATGCATCGCTG 1126  
QY 61 TGAATTTAAGCAACAGCTGTATAAAGCAAGCTCCACTGCAAGAGGNGGCCGCGGCAGGA 120  
Db 1127 TGAATTTAAGCAACAGCTGTATAAAGCAAGCTCCACTGCAAGAGGNGGCCGCGGCAGGA 1186  
QY 121 GAATCTCGCTTGTCCAAGACAGGGGCTTAAGGAGGGTCTCCACACTGCTGTAGGGCT 180  
Db 1187 GAATCTCGCTTGTCCAAGACAGGGGCTTAAGGAGGGTCTCCACACTGCTGTAGGGCT 1246  
QY 181 GTTGCAATTTTATTATTAGTAGAAAGTGGAAAGGCTCTTCTCAACTTTTTCCTTTGGGC 240  
Db 1247 GTTGCAATTTTATTATTAGTAGAAAGTGGAAAGGCTCTTCTCAACTTTTTCCTTTGGGC 1306  
QY 241 TGGAGAAATTTAGAATCAGAAAGTTTCTGGAGTTTTCAGGCTATCATATATATCTATCT 300  
Db 1307 TGGAGAAATTTAGAATCAGAAAGTTTCTGGAGTTTTCAGGCTATCATATATATCTATCT 1366  
QY 301 GAAAGGCAACATAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360  
Db 1367 GAAAGGCAACATAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1426  
QY 361 TTACTCCTAAAGGCTTCAATTTTAGTCCAGATTTTATAGTCTGGCTGCACTTAACCTATG 420  
Db 1427 TTACTCCTAAAGGCTTCAATTTTAGTCCAGATTTTATAGTCTGGCTGCACTTAACCTATG 1486  
QY 421 CCTCGCTTATTAGCCCGAGATCTGGTCTTTTNTGNTNTNTNTNTNTNTNTNTNTNTNTNT 480  
Db 1487 CCTCGCTTATTAGCCCGAGATCTGGTCTTTTNT - TTTTNTNTNTNTNTNTNTNTNTNT 1544  
QY 481 CAAAGCTTTATCTGTCTTGACCTTTTAAAGTTTGGGGCAGATTCGAATTTGGGCTA 540  
Db 1545 CAAAGCTTTATCTGTCTTGACCTTTTAAAGTTTGGGGCAGATTCGAATTTGGGCTA 1603  
QY 541 AAAGACATGCATTTTAAAGCTTAGCAACTCTTATTTCTCTCTCTCTCTCTCTCTCTCTCT 600  
Db 1604 AAAGACATGCATTTTAAAGCTTAGCAACTCTTATTTCTCTCTCTCTCTCTCTCTCTCTCT 1661  
QY 601 ATTAATCCCAATCTTATTAAAGACCTGACAGCTTGAGAAAGTCACTACTGCAATTTAT 660  
Db 1662 ATTAATCCCAATCTTATTAAAGACCTGACAGCTTGAGAAAGTCACTACTGCAATTTAT 1721  
QY 661 AGGACCTTCTGGTGTCTGCTGTTTGAAGTCTGACATCTCTTGAGAAATCTTTGCT 720  
Db 1722 AGGACCTTCTGGTGTCTGCTGTTTGAAGTCTGACATCTCTTGAGAAATCTTTGCT 1781  
QY 721 ATGACAGAGGTAAGAGGTAATGATTTTACAGAGGAAGAACACAGCGCAAGATGAAG 780  
Db 1782 ATGACAGAGGTAAGAGGTAATGATTTTACAGAGGAAGAACACAGCGCAAGATGAAG 1841  
QY 781 GGCCAGGCTTACTGAGGCTGTCCAGTGGAGGCTCATGGTGGGACATCGGAAAAAGAGGC 840

```
Db 1842 GGCACAGGCTTACTGA-GCTGTCAGTGGAGGCTCATGGGTGGGACATGAAAGAGGC 1900
Qy 841 AGCCTAGGCTTGGGAGCCGACGTCCTAGCAGCAAGGAGCTAGTGGAGCTTTTGC 900
Db 1901 AGCCTAGGCTTGGGAGCCGACGTCCTAGCAGCAAGGAGCTAGTGGAGCTTTTGC 1960
Qy 901 AGGAAAGGCTTAAGAAAGGAAAGCAATCTTAAACACACACAGAACTGTCCAAATGC 960
Db 1961 AGGAAAGGCTTAAGAAAGGAAAGCAATCTTAAACACACACAGAACTGTCCAAATGC 2020
Qy 961 TTTGGGAATCTGTTTATGCTTATAATGGGTCCCCAAATGGGTAACTAGACTTCAGA 1020
Db 2021 TTTGGGAATCTGTTTATGCTTATAATGGGTCCCCAAATGGGTAACTAGACTTCAGA 2080
Qy 1021 GAGAAATGAGCAGAGAGAGAGAGAGAAATCTGGCTGCTCCATTTTCACTTGTATCT 1080
Db 2081 GAGAAATGAGCAGAGAGAGAGAGAAATCTGGCTGCTCCATTTTCACTTGTATCT 2140
Qy 1081 CAGGTGAGCTGGTAGAGGGAGAGCATTAGAAAGAAATGAAACACAAACAACTACTAAT 1140
Db 2141 CAGGTGAGCTGGTAGAGGGAGAGCATTAGAAAGAAATGAAACACAAACAACTACTAAT 2200
Qy 1141 GAGTACGCTGAGGCTGGGAGTCTCTGACTCCACTACTTAATTCGGTTAGTGAGAAA 1200
Db 2201 GAGTACGCTGAGGCTGGGAGTCTCTGACTCCACTACTTAATTCGGTTAGTGAGAAA 2260
Qy 1201 CCTTTCATTTTCTTTTATGAGAGGCGCAGCTTACTGTTGGTGCAAAATGGCCACAT 1260
Db 2261 CCTTTCATTTTCTTTTATGAGAGGCGCAGCTTACTGTTGGTGCAAAATGGCCACAT 2320
Qy 1261 AAGTAAATAGAAAGTTGGCCAAATTCACCCCAATTTCTGGTTGGGCTCCACATTGCA 1320
Db 2321 AAGTAAATAGAAAGTTGGCCAAATTCACCCCAATTTCTGGTTGGGCTCCACATTGCA 2380
Qy 1321 ATGTTCAATGCCAGTGTCTGTGACACGACCGGAGTACTAGCCAGACAAAGGAGGG 1380
Db 2381 ATGTTCAATGCCAGTGTCTGTGACACGACCGGAGTACTAGCCAGACAAAGGAGGG 2440
Qy 1381 TAGCCTGAATTCCTTCTGCTCTTACATTTCTTTTAAATTAAGCATTTAGTCTCAGTC 1440
Db 2441 TAGCCTGAATTCCTTCTGCTCTTACATTTCTTTTAAATTAAGCATTTAGTCTCAGTC 2500
Qy 1441 CTAAGTACTGCTCTTCTCTCCCTCTCTGAAATTTAAATTAAGCATTTAGTCTCAGTC 1500
Db 2501 CTAAGTACTGCTCTTCTCTCCCTCTCTGAAATTTAAATTTCTTCAACTTGAATTTGC 1560
Qy 1501 AAGGATTAACATTTCTGCTGATGATATATGTTGACGAGGAAAGAAAGTGTCTT 1620
Db 2561 AAGGATTAACATTTCTGCTGATGATATATGTTGACGAGGAAAGAAAGTGTCTT 2619
Qy 1561 TGTGTTAAATTAATGTTGTTGTAATCCATCTTCTGTTTCCCATTTGGAAGTGTCTT 1620
Db 2620 TGTGTTAAATTAATGTTGTTGTAATCCATCTTCTGTTTCCCATTTGGAAGTGTCTT 2679
Qy 1621 AACCCTCTCTGACTGGTGAAGAAACATCTGAAGAGTGTCTATCAGCATCTGCACAGG 1680
Db 2680 AACCCTCTCTGACTGGTGAAGAAACATCTGAAGAGTGTCTATCAGCATCTGCACAGG 2739
Qy 1681 TGAATGGATGGTCTCAGAACCATTTTCCACGACAGCCTGTTTCTATCTCTGTTTAAAT 1740
Db 2740 TGAATGGATGGTCTCAGAACCATTTTCCACGACAGCCTGTTTCTATCTCTGTTTAAAT 2799
Qy 1741 AATTAGTTGGTCTCTACATGATATAACAAACCTGCTCCAACTGTGCATATAAAGTC 1800
Db 2800 AATTAGTTGGTCTCTACATGATATAACAAACCTGCTCCAACTGTGCATATAAAGTC 2859
Qy 1801 TGTGACTTGAAGTTAGTCAGACACCCCAACAACTTTTATTTCTATGTTTGTGCA 1860
Db 2860 TGTGACTTGAAGTTAGTCAGACACCCCAACAACTTTTATTTCTATGTTTGTGCA 2919
Qy 1861 ACATATGAGTGTGTTGAAATTAAGTACCCATGCTTTTATTAAGAAAGGAAAGGCA 1916
Db 1916 ACATATGAGTGTGTTGAAATTAAGTACCCATGCTTTTATTAAGAAAGGAAAGGCA 1976
```

```
Db 2920 ACATATGAGTGTGTTGAAATTAAGTACCCATGCTTTTATTAAGAAAGGAAAGGCA 2975
RESULT 7
US-10-012-896-335
; Sequence 335, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Panger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Megher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 335
; LENGTH: 2984
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-335
Query Match 94.9%; Score 1836.4; DB 13; Length 2984;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1897; Conservative 0; Mismatches 12; Indels 7; Gaps 5;
```

```
Qy 1 AATGGTATGCCAACTTAAGTATTACAGGGTGGCCCAATAGAACAGATGCACTCGCTG 60
Db 1067 AATGGTATGCCAACTTAAGTATTACAGGGTGGCCCAATAGAACAGATGCACTCGCTG 1126
Qy 61 TGATTTTAAAGCAAGCTGTATATAACAGAACTCACTGCAAGAGGGGGCGGGCCAGGA 120
Db 1127 TGATTTTAAAGCAAGCTGTATATAACAGAACTCACTGCAAGAGGGGGCGGGCCAGGA 1186
Qy 121 GAATCTCCGCTTGTCCAGACAGGGGCTTAGAGGGTCTCCACACTGCTCTAGGGGCT 180
Db 1187 GAATCTCCGCTTGTCCAGACAGGGGCTTAGAGGGTCTCCACACTGCTCTAGGGGCT 1246
Qy 181 GTTGCAATTTTATTATAGTAAAGAGGCTCTTCTCAACTTTTCCCTTGGGCT 240
Db 1247 GTTGCAATTTTATTATAGTAAAGAGGCTCTTCTCAACTTTTCCCTTGGGCT 1306
Qy 241 TGGAGAAATTTAGAAATCAGAACTTCTCGAGTCTTTCAGGCTATCATATATCTATCT 300
Db 1307 TGGAGAAATTTAGAAATCAGAACTTCTCGAGTCTTTCAGGCTATCATATATCTATCT 1366
Qy 301 GAAAGGCAACATAATCTTCTCCCTCTTTTAAATTTTGTCTCTTTTTCGAGCA 360
Db 1367 GAAAGGCAACATAATCTTCTCCCTCTTTTAAATTTTGTCTCTTTTTCGAGCA 1426
```



QY 121 GAATCTCCGCTGTTCAGAGACAGGGGCTTAAGAGGGTCTCCACATGCTGTAGGGCT 180  
Db |||||  
QY 1187 GAATCTCCGCTGTTCAGAGACAGGGGCTTAAGAGGGTCTCCACATGCTGTAGGGCT 1246  
Db |||||  
QY 181 GTTGCAATTTTATTAGTAGAAGAGTGAAGGCTCTTCTCAACTTTTTTCCCTTGGGC 240  
Db |||||  
QY 1247 GTTGCAATTTTATTAGTAGAAGAGTGAAGGCTCTTCTCAACTTTTTTCCCTTGGGC 1306  
Db |||||  
QY 241 TGGAGAAATTTAGAAATCAGAAATTTTCCCTGGAGTTTCCAGGCTATCATATATCTGTATCT 300  
Db |||||  
QY 1307 TGGAGAAATTTAGAAATCAGAAATTTTCCCTGGAGTTTCCAGGCTATCATATATCTGTATCT 1366  
Db |||||  
QY 301 GAAAGGCAACATTAATCTCTCCCTCCCTTTTAAATTTTGTGTCCTTTTGGAGAA 360  
Db |||||  
QY 1367 GAAAGGCAACATTAATCTCTCCCTCCCTTTTAAATTTTGTGTCCTTTTGGAGAA 1426  
Db |||||  
QY 361 TTACTCTAAAGGGCTTCAATTTAGTCCAGATTTTATGTCGGCTGCACCTAACTTATG 420  
Db |||||  
QY 1427 TTACTCTAAAGGGCTTCAATTTAGTCCAGATTTTATGTCGGCTGCACCTAACTTATG 1486  
Db |||||  
QY 421 CCTCGCTTATTAGCCGAGATCTGGTCTTTTTTNGTNTTTTTTTTTTTTTTTTTTTTT 480  
Db |||||  
QY 1487 CCTCGCTTATTAGCCGAGATCTGGTCTTTTTT - TTTTTTTTTTTTTTCCGTCCTCC 1544  
Db |||||  
QY 481 CAAAGCTTTATCTGCTGACCTTTTAAAGAGTTTGGGGCAGATCTGAATTTGGGCTA 540  
Db |||||  
QY 1545 CAAAGCTTTATCTGCTGACCTTTTAAAGAGTTTGGGGCAGATCTGAATTTGGCTA 1603  
Db |||||  
QY 541 AAGACATGCAATTTTAAAGCTAGGCACTCTTAATTTCTTTTAAAGAGTTTGGGGCAGATCTGAATTTGGGCTA 600  
Db |||||  
QY 1604 AAGACATGCAATTTTAAAGCTAGGCACTCTTAATTTCTTTTAAAGAGTTTGGGGCAGATCTGAATTTGGGCTA 1661  
Db |||||  
QY 601 ATTAAATCCCAATCTTATTAAAGCTAGGCACTCTTAATTTCTTTTAAAGAGTTTGGGGCAGATCTGAATTTGGGCTA 660  
Db |||||  
QY 1662 ATTAAATCCCAATCTTATTAAAGCTAGGCACTCTTAATTTCTTTTAAAGAGTTTGGGGCAGATCTGAATTTGGGCTA 1721  
Db |||||  
QY 661 AGGACCTTCTGGTGTCTGCTGTGTTGAGTCTGCAATCTTAAAGAGTTTGGGGCAGATCTGAATTTGGGCTA 720  
Db |||||  
QY 1722 AGGACCTTCTGGTGTCTGCTGTGTTGAGTCTGCAATCTTAAAGAGTTTGGGGCAGATCTGAATTTGGGCTA 1781  
Db |||||  
QY 721 ATGCAGAGGCTAAGAGTATGATTTTCAAGAGGAGAAACACAGCGCAGAAATGAAG 780  
Db |||||  
QY 1782 ATGCAGAGGCTAAGAGTATGATTTTCAAGAGGAGAAACACAGCGCAGAAATGAAG 1841  
Db |||||  
QY 781 GCCCAGGCTTACTGAGGCTGTCAGTGGAGGCTGAGTGGAGGCTGAGTGGAGGCTTGGC 900  
Db |||||  
QY 1842 GCCCAGGCTTACTGA - GCTGTCAGTGGAGGCTGAGTGGAGGCTGAGTGGAGGCTTGGC 960  
Db |||||  
QY 841 AGCTAGGCTTGGAGGCTGAGTGGAGGCTGAGTGGAGGCTGAGTGGAGGCTTGGC 1900  
Db |||||  
QY 1901 AGCTAGGCTTGGAGGCTGAGTGGAGGCTGAGTGGAGGCTGAGTGGAGGCTTGGC 1960  
Db |||||  
QY 901 AGGAAAGGCTTAAAGAAAGGAAACCATTTTAAAGACACACAGAAATCTGCAAAATGC 960  
Db |||||  
QY 1961 AGGAAAGGCTTAAAGAAAGGAAACCATTTTAAAGACACACAGAAATCTGCAAAATGC 2020  
Db |||||  
QY 961 TTGGGAACTGTGTTTATTCCTTATTAATGGGTCCTTAAAGTGGTAACTGAGCTTCA 1020  
Db |||||  
QY 2021 TTGGGAACTGTGTTTATTCCTTATTAATGGGTCCTTAAAGTGGTAACTGAGCTTCA 2080  
Db |||||  
QY 1021 GAGAAATGAGCAGAGCAGAGGAGAAATCTGGTGTCTTCCATTTTCAATTTCTGTATCT 1080  
Db |||||  
QY 2081 GAGAAATGAGCAGAGCAGAGGAGAAATCTGGTGTCTTCCATTTTCAATTTCTGTATCT 2140  
Db |||||  
QY 1081 CAGTGTGCTGTAGGGGAGACATTTAGAAAAAATGAAACAAACAAACAAATTAAT 1140  
Db |||||  
QY 2141 CAGTGTGCTGTAGGGGAGACATTTAGAAAAAATGAAACAAACAAACAAATTAAT 2200  
Db |||||  
QY 1141 GAGTGTGCTGTAGGGGAGACATTTGACTTCACTTAAATTCCTTTAGTGAGAA 1200  
Db |||||  
QY 2201 GAGTGTGCTGTAGGGGAGACATTTGACTTCACTTAAATTCCTTTAGTGAGAA 2260  
Db |||||  
QY 1201 CCTTCAATTTCTTTTATTAGAGGGCAGCTTACTGTGTGGCAAAATTTGCAACAT 1260  
Db |||||

RESULT 9

US-10-205-823-282  
; Sequence 282, Application US/10205823  
; Publication No. US20030108963A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Gannavarapu, Manjula  
; APPLICANT: Gorbacheva, Bella  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Wensey, Angela M.  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Anderson, Dustin  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER  
; FILE REFERENCE: MRI-044  
; CURRENT APPLICATION NUMBER: US/10/205,823  
; CURRENT FILING DATE: 2002-07-25  
; PRIOR APPLICATION NUMBER: 60/307,982  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: 60/314,356

```
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282
; LENGTH: 3266
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-205-823-282

Query Match      85.0%; Score 1645.6; DB 14; Length 3266;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1826; Conservative 0; Mismatches 47; Indels 34; Gaps 12;

QY 1 AATGGTATGCCAATTAAGTATTTACAGGTGGCCCAATAGAACAAAGATGCATCGCTG 60
DB 1373 AATGGTATGCCAATTAAGTATTTACAGGTGGCCCAATAGAACAAAGATGCATCGCTG 1432
QY 61 TGAATTTTAACACAGCTGTATTAACAGAACTCCACTGCAAGAGGGGGGGCCGAGGA 120
DB 1433 TGAATTTTAACACAGCTGTATTAACAGAACTCCACTGCAAGAGGGGGGGCCGAGGA 1492
QY 121 GAATCTCCGCTGTGTCGAAGACAGGGGCTTAAGGAGGCTCTCCACACTGCTGTAGGGCT 180
DB 1493 GAATCTCCGCTGTGTCGAAGACAGGGGCTTAAGGAGGCTCTCCACACTGCTGTAGGGCT 1552
QY 181 GTTGCAATTTTTTATTAGTAGAAGAGTGAAGGCTCTCTCAACTTTTTTCCCTTGGGC 240
DB 1553 GTTGCAATTTTTTATTAGTAGAAGAGTGAAGGCTCTCTCAACTTTTTTCCCTTGGGC 1612
QY 241 TGGAGAAATTTAGATCAGAGTTTCCTGGAGTTTTCAGGCTATCATATATACGTATCCT 300
DB 1613 TGGAGAAATTTAGATCAGAGTTTCCTGGAGTTTTCAGGCTATCATATATACGTATCCT 1672
QY 301 GAAAGGCAACATAATCTCTCCCTCCCTTTTAAATTTTGTGTTCTTTTTCGAGCAA 360
DB 1673 GAAAGGCAACATAATCTCTCTCCCTCCCTTTTAAATTTTGTGTTCTTTTTCGAGCAA 1732
QY 361 TTAATCACTAAGGGCTTCAATTTAGTCCAGATTTTGTGCTGGCTGCACCTAACTATG 420
DB 1733 TTAATCACTAAGGGCTTCAATTTAGTCCAGATTTTGTGCTGGCTGCACCTAACTATG 1792
QY 421 CCTCGCTATTATAGCCGAGATCTGGTCTTTTNTGTTNTTTTNTTTCGCTCCC 480
DB 1793 CCTCGCTATTATAGCCGAGATCCGGTCTTTTNTGTTNTTTTNTTTCGCTCCC 1848
QY 481 CAAAGCTTTATCTGTCTTGACTTTTAAAGAGTTTGGGGGAGATTTCTGAATTTGGGCTA 540
DB 1849 CAAAGCTTTATCTGTCTTGACTTTTAAAGAGTTTGGGGGAGATTTCTGAATTTGGGCTA 1907
QY 541 AAAGACATGATTTTAAATCTAGGCAACTCTTATTTCTTTTCTTTTAAATATACATAGC 600
DB 1908 AAAGACATGATTTTAAATCTAGGCAACTCTTATTTCTTTTCTTTTAAATATACATAGC 1965
QY 601 ATTAATCCCAATCTTATTTAAAGACCTGACAGCTTGAGAA--GGTCACTACTGCAATTTA 659
DB 1966 ATTAATCCCAATCTTATTTAAAGACCTGACAGCTTTGAGAGGGTCACTACTGCAATTTA 2025
QY 660 TAGGACCTTCTGTGGTTCGTGTACGTTTGAAGTCTGACAACTCTTGAGAAATCTTTTG 719
DB 2026 TAGGACCTTCTGTGGTTCGTGTACGTTTGAAGTCTGACAACTCTTGAGAAATCTTTTG 2085
QY 720 CATGCAGAGGAGTAAAGAGTATTTGATTTTTCAGAG--GGAAGAACACAGCGCAGAAATGA 778
DB 2086 CATGCAGAGGAGTAAAGAGTATTTGATTTTTCAGAGGAGGAGAACACAGCGCAGAAATGA 2145
QY 779 AGGCCAGGCTTACTAGGCTGCTCAGTGGAGGCTCATGGTGGAGCATGGGAAAGAG 838
DB 779 AGGCCAGGCTTACTAGGCTGCTCAGTGGAGGCTCATGGTGGAGCATGGGAAAGAG 838
```

```
DB 2146 AGGCGCAGGCTTACTGA--CCTGTCCAGTGGAGGGGCTCATGGTGGGACATGGAAGAAG 2204
QY 839 GCAGCCTAGGCCCTGGGGAGCCAGTCCACTGAGCAAGCAAGGAGCTGAGTGAAGCCTTTT 898
DB 2205 GCAGCCTAGGCCCTGGGGAGCCAGTCCACTGAGCAAGCAAGGAGCTGAGTGAAGCCTTTT 2264
QY 899 GCAGGAAAAGGCTAGAAAAGGAAAACCATTTCTTAAACACAAAGAACTGTTCAAAAT 958
DB 2265 GCAGGAAAAGGCTAGAAAAGGAAAACCATTTCTTAAACACAAAGAACTGTTCAAAAT 2324
QY 959 GCTTTGGGAACCTGTGTTTATTGCTTATAAT--GGGTCCCAAAATGGTAACTTAGACTTC 1017
DB 2325 GCTTTGGGAACCTGTGTTTATTGCTTATAATGGGGTCCCCAAATGGTAACTTAGACTTC 2384
QY 1018 AGAGAAATGAGCAGAGAGCAAGAGGAAATCT--GGCTGTCTCTTCCATTTTCAATCTGT 1076
DB 2385 AGAGAAATGAGCAGAGAGCAAGAGGAAATCTGGGCTGTCTCTTCCATTTTCAATCTGT 2444
QY 1077 ATCTCAGGTGAGCTGGTAGAGGGGAGACATTTAGAAAAAATGAAACAAACAAATTTAC 1136
DB 2445 ATCTCAGGTGAGCTGGTAGAGGGGAGACATTTAGAAAAAATGAAACAAACAAATTTAC 2504
QY 1137 TAATGAGGTACGCTGAGGCTCGGAGTCTCTT-----GACTCCACTACTTAATTC 1187
DB 2505 TAATGAGGTACGCTGAGGCTCGGAGTCTCTTGAGCTCCAGCACTTCAAAATTTAAATGA 2564
QY 1188 GTTTAGTAGAAGAACCTTTTCAATTTTCTTTTATTAGAAAGGCCAG-----CTTACTGT 1239
DB 2565 GCCATGAGTCAAAACCACTGCAATCCAGCCTGGGCAACGAGCAAGACCCAGTCTTACTGT 2624
QY 1240 TGGTGGCAAAATTTGCCAACAATTAAGTTAATAGAAAAGTTGGCCAAATTTCAACCCATTTCTG 1299
DB 2625 TGGTGGCAAAATTTGCCAACAATTAAGTTAATAGAAAAGTTGGCCAAATTTCAACCCATTTCTG 2684
QY 1300 TGGTTTGGGCTCCACATTTGCAATGTTCAATGCCAGCTGCTGACCGACCGCGAGTAC 1359
DB 2685 TGGTTTGGGCTCCACATTTGCAATGTTCAATGCCAGCTGCTGACCGACCGCGAGTAC 2744
QY 1360 TAGCCAGCACAAAAGGCGAGGTAAGTGAATTTGCTTTTCTGCTCTTTTACATTTCTTTTAAA 1419
DB 2745 TAGCCAGCACAAAAGGCGAGGTAAGTGAATTTGCTTTTCTGCTCTTTTACATTTCTTTTAAA 2804
QY 1420 ATAAGCATTTAGTGTCTAGTCCCTACTGAGTACTCTTTCTCTCCCTCTCTCTGAATTTAA 1479
DB 2805 ATAAGCATTTAGTGTCTAGTCCCTACTGAGTACTCTTTCTCTCCCTCTCTCTGAATTTAA 2864
QY 1480 TTCTTTCACTTTGCAATTTGCAAGGATTTACACATTTCACTGCTGATGATATTTGTGTGCA 1539
DB 2865 TTCTTTCACTTTGCAATTTGCAAGGATTTACACATTTCACTGCTGATGATATTTGTGTGCA 2924
QY 1540 GNGAAAAGAAAAAGTGTCTTTTGTAAATTTACTTTGTTTGTGAATCCATCTTGTCTTTT 1599
DB 2925 ---AAAAAAAAGAGTGTCTTTTGTAAATTTACTTTGTTTGTGAATCCATCTTGTG--TT 2979
QY 1600 TCCCATTTGGAACCTAGTCAATTAACCCACTCTGAACTGGTAGAAAAACATCTGAAGAGCT 1659
DB 2980 TCCCATTTGGAACCTAGTCAATTAACCCACTCTCTGAACTGGTAGAAAAACATCTGAAGAGCT 3039
QY 1660 AGTCTATCAGCATCTGACAGGTGAATTTGATGGTCTCTGAGAACCATTTTCAACCCAGCAGC 1719
DB 3040 AGTCTATCAGCATCTGACAGGTGAATTTGATGGTCTCTGAGAACCATTTTCAACCCAGCAGC 3099
QY 1720 CTGTTTCTATCTGTTTAAATAATTTAGTTTGGGTTCTCTTACATGATATACAAACCTGTCT 1779
DB 3100 CTGTTTCTATCTGTTTAAATAATTTAGTTTGGGTTCTCTTACATGATATACAAACCTGTCT 3159
QY 1780 CCAATCTGTACATAAAAGTCTGTGACTTTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 1839
DB 3160 CCAATCTGTACATAAAAGTCTGTGACTTTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 3219
QY 1840 TTTTCTATGTGTTTGTGCAACATATGAGTGTGTTTGTGAAATTTAAAGT 1886
DB 3220 TTTTCTATGTGTTTGTGCAACATATGAGTGTGTTTGTGAAATTTAAAGT 3266
```

RESULT 10  
 US-10-202-193-334/c  
 ; Sequence 334, Application US/10202193  
 ; Publication No. US20020192699A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Zhang, Jimmy  
 ; APPLICANT: Astel, Jon H.  
 ; APPLICANT: Carroll III, Eddie  
 ; APPLICANT: Endege, Wilson O.  
 ; APPLICANT: Ford, Donna M.  
 ; APPLICANT: Monahan, John E.  
 ; APPLICANT: Schlegel, Robert  
 ; APPLICANT: Steinmann, Kathleen E.  
 ; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT  
 ; FILE REFERENCE: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER  
 ; CURRENT APPLICATION NUMBER: PP-01532.103/200130.463D1  
 ; CURRENT FILING DATE: 2002-07-23  
 ; NUMBER OF SEQ ID NOS: 341  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 334  
 ; LENGTH: 2051  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1) - (2051)  
 ; OTHER INFORMATION: n = A, T, C or G  
 US-10-202-193-334

Query Match 71.7%; Score 1388.8; DB 13; Length 2051;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 1438; Conservative 1; Mismatches 12; Indels 5; Gaps 4;

Qy	449	TTTTTTTGTGTTTTTTTTTTTTTTTCCGCTCCCAAGCTTTATCTGCTTGCACCTTTTAA	508
Db	1814	TGTTTTTTTTTTTTTTTTTTTTTTCGGCTCCCAAGCTTTATCTGCTTGCACCTTTTAA	1755
Qy	509	AAAAGTTGGGGGAGAGTCTGAAATGGGCTAAAGACATGCTATTTTAAACATAGCA	568
Db	1754	AAAAGTTGGGGGAGAGTCTGAAATGGGCTAAAGACATGCTATTTTAAACATAGCA	568
Qy	569	CTTCTTATTTCTTCTTTTAAATAATAGATTAATCCCAATCCTATTTTAAAGACC	628
Db	1697	ACTCTATTTCTTCTTTTAAATAATAGATTAATCCCAATCCTATTTTAAAGACC	1638
Qy	629	TGACAGCTTGAGAGGTCTACTGCAATTTATAGACCTTCTGGTGGTTCTGCTGTAGC	688
Db	1637	TGACAGCTTGAGAGGTCTACTGCAATTTATAGACCTTCTGGTGGTTCTGCTGTAGC	1578
Qy	689	TTTGAAGTCTGCAATCTTGAGAACTTTTGCAATGAGAGGTAAAGAGTATTGGATT	748
Db	1577	TTTGAAGTCTGCAATCTTGAGAACTTTTGCAATGAGAGGTAAAGAGTATTGGATT	1518
Qy	749	TTGACAGAGGAAGAACACAGCGCAGAAATGAAGGCGCAGGCTTACTGAGCTGTCCAGTGG	808
Db	1517	TTGACAGAGGAAGAACACAGCGCAGAAATGAAGGCGCAGGCTTACTGAGCTGTCCAGTGG	1459
Qy	809	AGGCTCATGCTGGTGAGACATGGAAGAGGAGGAGGCTAGGCCCTGGGAGCCAGTCCAC	868
Db	1458	AGGCTCATGCTGGTGAGACATGGAAGAGGAGGAGGCTAGGCCCTGGGAGCCAGTCCAC	1399
Qy	869	TGACAGAGGAAGGAATGAGTGGAGCTTTTGCAAGAAAGGCTTAAGAAAGGAACCA	928
Db	1398	TGACAGAGGAAGGAATGAGTGGAGCTTTTGCAAGAAAGGCTTAAGAAAGGAACCA	1339
Qy	929	TTCTAAACACACAGAGAACTGTCCTTGGGAACTGTTTATTTGCTTATAAT	988
Db	1338	TTCTAAACACACAGAGAACTGTCCTTGGGAACTGTTTATTTGCTTATAAT	1279
Qy	989	GGGTCCCCAAATGGGTAACTTAGACTTCAGAGAGATGACAGAGCAAGAGAAAT	1048

Db	1278	GGGTCCCCAAATGGGTAACTTAGACTTCAGAGAGATGACAGAGCAAGAGAAAT	1219
Qy	1049	CTGGCTGTCCTTCCATTTTCAATCTGTATCTCAGGTGAGCTGTAGAGGGAGACATTA	1108
Db	1218	CTGGCTGTCCTTCCATTTTCAATCTGTATCTCAGGTGAGCTGTAGAGGGAGACATTA	1159
Qy	1109	GAAGAAATGAAGCAACAAACAAATTAATAGAGTACGCTGAGCCCTGGAGTCTCTT	1168
Db	1158	GAAGAAATGAAGCAACAAACAAATTAATAGAGTACGCTGAGCCCTGGAGTCTCTT	1099
Qy	1169	GACTCCACTACTTAATTCGGTTTAGTGAGAAACCTTTCAATTTCTTTTATAGAGGC	1228
Db	1098	GACTCCACTACTTAATTCGGTTTAGTGAGAAACCTTTCAATTTCTTTTATAGAGGC	1039
Qy	1229	CAGCTTACTGTTGGTGCAAAATTCGCAACATAAGTAAATAGAAAGTTGGCAATTCAC	1288
Db	1038	CAGCTTACTGTTGGTGCAAAATTCGCAACATAAGTAAATAGAAAGTTGGCAATTCAC	979
Qy	1289	CCCATTTTCTGTGGTTTGGGCTCCACATTCGAATGTTCAATGCCACGCTGCTGACACC	1348
Db	978	CCCATTTTCTGTGGTTTGGGCTCCACATTCGAATGTTCAATGCCACGCTGCTGACACC	919
Qy	1349	GACCGAGTACTAGCCAGCAACAAAGCAGGCTAGCTGAATTCCTTCTGCTCTTTACA	1408
Db	918	GACCGAGTACTAGCCAGCAACAAAGCAGGCTAGCTGAATTCCTTCTGCTCTTTACA	859
Qy	1409	TTTCTTTTAAATAAGCAATTTAGTCTGCTCCCTACTGAGTACTCTTCTCTCCCTCC	1468
Db	858	TTTCTTTTAAATAAGCAATTTAGTCTGCTCCCTACTGAGTACTCTTCTCTCCCTCC	799
Qy	1469	TCTGAATTAATCTTTCAACTTTCGAATTTGCAAGATTAACATTTCTGCTGATGAT	1528
Db	798	TCTGAATTAATCTTTCAACTTTCGAATTTGCAAGATTAACATTTCTGCTGATGAT	739
Qy	1529	ATTGCTGTCAGAGAAAGAAAGTCTTTTGTAAATTAATTTGCTGTTGTAATCC	1588
Db	738	ATTGCTGTCAGAGAAAGAAAGTCTTTTGTAAATTAATTTGCTGTTGTAATCC	680
Qy	1589	ATCTGCTTTTCCCACTGGAACTAGTCAATTAACCTCTCTGAACTGTTGTAATCC	1648
Db	679	ATCTGCTTTTCCCACTGGAACTAGTCAATTAACCTCTCTGAACTGTTGTAATCC	620
Qy	1649	TCTGAAGAGTACTGTCATCAGCATCTGACAGTGAATTTGATGTTCTCAGAACCAATTC	1708
Db	619	TCTGAAGAGTACTGTCATCAGCATCTGACAGTGAATTTGATGTTCTCAGAACCAATTC	560
Qy	1709	ACCAGACAGCTGTTTCTATCTGTTTAAATAATTAATTTGCTGTTGTAATCC	1768
Db	559	ACCAGACAGCTGTTTCTATCTGTTTAAATAATTAATTTGCTGTTGTAATCC	500
Qy	1769	CAAAACCTGCTCCAATCTGTCACATTAAGTCTGCTGAAAGTTAGTCAGCACCCCC	1828
Db	499	CAAAACCTGCTCCAATCTGTCACATTAAGTCTGCTGAAAGTTAGTCAGCACCCCC	440
Qy	1829	ACCAACCTTATTTTCTATGTTTTCACATATAGTCTTTTGAAGTAAAGTAC	1888
Db	439	ACCAACCTTATTTTCTATGTTTTCACATATAGTCTTTTGAAGTAAAGTAC	380
Qy	1889	CCATGCTTTTATTTAA 1904	
Db	379	CCATGCTTTTATTTAA 364	

RESULT 11  
 US-09-925-300-258  
 ; Sequence 258, Application US/09925300  
 ; Patent No. US20020151681A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Craig Rosen,  
 ; APPLICANT: Steve Ruben,  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA101

; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 258  
; LENGTH: 755  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (755)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-300-258

Query Match 35.4%; Score 686.2; DB 10; Length 755;  
Best Local Similarity 98.9%; Pred. No. 2.1e-176;  
Matches 700; Conservative 1; Mismatches 5; Indels 2; Gaps 1;  
QY 1209 TTTTCTTTTATTAGAGGGCCAGCTTACTGTTGGTGGCAAAATTGCGCAATAGTTAAT 1268  
Db |||||  
QY 38 TTTTCTTTTWTAGAGGGCCAGCTTACTGTTGGTGGCAAAATTGCGCAATAGTTAAT 97  
Db |||||  
QY 1269 AGAAGTGGCCAAATTCACCCCAATTTCTGTGTTTGGGCTCCACATGCAATGTTCAA 1328  
Db |||||  
QY 98 AGAAGTGGCCAAATTCACCCCAATTTCTGTGTTTGGGCTCCACATGCAATGTTCAA 157  
Db |||||  
QY 1329 TGGCAGTGTGCTGACACCGAGGAGTACTGACGAGCAACAAAGGCGAGGTAGCTGA 1388  
Db |||||  
QY 158 TGGCAGTGTGCTGACACCGAGGAGTACTGACGAGCAACAAAGGCGAGGTAGCTGA 217  
Db |||||  
QY 1389 ATTGCTTTCTGCTCTTACATTTCTTTTAAATTAAGCATTTAGTGTCTCAGTCCCTACTGA 1448  
Db |||||  
QY 218 ATTGCTTTCTGCTCTTACATTTCTTTTAAATTAAGCATTTAGTGTCTCAGTCCCTACTGA 277  
Db |||||  
QY 1449 GTACTCTTCTCTCCCTCTCTGAAATTTAAATTTCTTCAACTTGCATTTGCAAGGATTA 1508  
Db |||||  
QY 278 GTACTCTTCTCTCCCTCTCTGAAATTTAAATTTCTTCAACTTGCATTTGCAAGGATTA 337  
Db |||||  
QY 1509 CACATTTCTGATGATATATTTGTTGTCGAGNAGAAAGAAAGTGCTTTGTTTAA 1568  
Db |||||  
QY 338 CACATTTCTGATGATATATTTGTTGTCGA--AAAAAAGAAAGTGCTTTGTTTAA 395  
Db |||||  
QY 1569 ATTACTTGGTTGTGATCATCTGCTTTTCCCAATTTGGAACTAGTCAATTAACCCATC 1628  
Db |||||  
QY 396 ATTACTTGGTTGTGATCATCTGCTTTTCCCAATTTGGAACTAGTCAATTAACCCATC 455  
Db |||||  
QY 1629 TCTGAATGGTAGAAAAACATCTGAAGAGCTAGTCTATCAGCATCTGACAGGTGAATTGG 1688  
Db |||||  
QY 456 TCTGAATGGTAGAAAAACATCTGAAGAGCTAGTCTATCAGCATCTGACAGGTGAATTGG 515  
Db |||||  
QY 1689 ATGGTTCTCAGAACCAATTCACCCAGACAGCTGTTTCTATCTGTTTAAATTAATAGTT 1748  
Db |||||  
QY 516 ATGGTTCTCAGAACCAATTCACCCAGACAGCTGTTTCTATCTGTTTAAATTAATAGTT 575  
Db |||||  
QY 1749 TGGGTTCTCTACATGATACAAACCTGCTCCAACTCTGTCATATAAAGTCTGTGACTT 1808  
Db |||||  
QY 576 TGGGTTCTCTACATGATACAAACCTGCTCCAACTCTGTCATATAAAGTCTGTGACTT 635  
Db |||||  
QY 1809 GAAGTTTGTAGTACGACCCCAACCAACTTTATTTTCTATGTTTGTGTTTGTGCAACATATGA 1868  
Db |||||  
QY 636 GAAGTTTGTAGTACGACCCCAACCAACTTTATTTTCTATGTTTGTGTTTGTGCAACATATGA 695  
Db |||||  
QY 1869 GTGTTTGTGAAATTAAGTACCCATGCTTTATTTAAATAAATAAATAA 1916  
Db |||||  
QY 696 GTGTTTGTGAAATTAAGTACCCATGCTTTATTTAGAAAAAATAAATAA 743  
Db |||||

RESULT 12  
US-10-202-193-81/c

; Sequence 81, Application US/10202193  
; Publication No. US20020192699A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Jimmy  
; APPLICANT: Astel, Jon H.  
; APPLICANT: Carroll III, Eddie  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Ford, Donna M.  
; APPLICANT: Monahan, John E.  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Steinmann, Kathleen E.  
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT  
; FILE REFERENCE: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER  
; CURRENT APPLICATION NUMBER: US/10/202,193  
; CURRENT FILING DATE: 2002-07-23  
; NUMBER OF SEQ ID NOS: 341  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 81  
; LENGTH: 1024  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(1024)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-202-193-81  
Query Match 33.1%; Score 641.2; DB 13; Length 1024;  
Best Local Similarity 97.3%; Pred. No. 4.7e-164;  
Matches 681; Conservative 0; Mismatches 16; Indels 3; Gaps 3;  
QY 448 CTTTTTNTGNTTTTTTTTTTTTTTTTTCCTCCCTCCCAAGCTTTTATCTGTCTGACATTTTAA 507  
Db |||||  
QY 799 CTCCTTNNNNNNTTTTTTTTTTCNGTCTCCCTCCCAAGCTTTTATCTGTCTGACATTTTAA 740  
Db |||||  
QY 508 AAAAAATTTGGGGGAGATTTCTGAATTTGGGTAAAGAGCATGCAATTTTAAAGCATAGGCA 567  
Db |||||  
QY 739 AAAAAATTTGGGGGAGATTTCTGAATTTGGTAAAGAGCATGCAATTTTAAAGCATGCA 682  
Db |||||  
QY 568 ACTTCTTATTTCTTTCCTTTTAAATATACATAGCATTAATCCCAATCTTATTTAAAGAC 627  
Db |||||  
QY 681 ACTTCTTATTTCTTTCCTTTTAAATATACATAGCATTAATCCCAATCTTATTTAAAGAC 622  
Db |||||  
QY 628 CTGACAGCTTGAGAAAGGTCATCTACTGCAATTTATAGGACCTTCTGGTGGTCTGCTGTAC 687  
Db |||||  
QY 621 CTGACAGCTTGAGAAAGGTCATCTACTGCAATTTATAGGACCTTCTGGTGGTCTGCTGTAC 562  
Db |||||  
QY 688 GTTTGAAGTCTGACAAATCTTTGAGAAATCTTTGATGCAATGTCAGAGGAGGTAAGAGTATGGAT 747  
Db |||||  
QY 561 GTTTGAAGTCTGACAAATCTTTGAGAAATCTTTGATGCAATGTCAGAGGAGGTAAGAGTATGGAT 502  
Db |||||  
QY 748 TTTTCAGAGGAGAAACACAGCGCAGAAATGAGGGCCAGGCTTACTGAGGCTGTCAGTG 807  
Db |||||  
QY 501 TTTTCAGAGGAGAAACACAGCGCAGAAATGAGGGCCAGGCTTACTGAGGCTGTCAGTG 443  
Db |||||  
QY 808 GAGGCTCATGGTGGGACATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 867  
Db |||||  
QY 442 GAGGCTCATGGTGGGACATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 383  
Db |||||  
QY 868 CTGAGCAAGCAAGGAGTGTGAGGCTTTTTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 927  
Db |||||  
QY 382 CTGAGCAAGCAAGGAGTGTGAGGCTTTTTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 323  
Db |||||  
QY 928 ATTCTTAAACACAAAGAACTGTCCAAATGCTTTGGAACTGTGTTTATTTGCTATAA 987  
Db |||||  
QY 322 ATTCTTAAACACAAAGAACTGTCCAAATGCTTTGGAACTGTGTTTATTTGCTATAA 263  
Db |||||  
QY 988 TGGGTCCCAAAATGCGGTAACTCTAGACTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1047  
Db |||||  
QY 262 TGGGTCCCAAAATGCGGTAACTCTAGACTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 203  
Db |||||  
QY 1048 TCTGGCTGTCTTCCATTTTTCATTTCTCTAGGTGAGCTGCTGAGAGGAGAGAGATTT 1107  
Db |||||

...TGGGCTTGAGCCCTTTTCAGGATAAAGGC-TAAGAAAAAGGAAAACCATTC 320

722 TGCAGAGGAGGTAAGAGGTATTGGATTTCACAGAGGAAGAACACAGCGCAGAAATGAAGG 781

```
Db 372 TG CAGAGGAGGTAAAGGTAATTGGATTTTTCACAGAGGAAGAACACAGCGCAGAAATGAAGG 431
Qy 782 GC CAGAGGCTTACTGAGGCTGTCCAGTGGAGGGCTCATGGGTGGGACATGGAAAAAGAGGCA 841
Db 432 GC CAGAGGCTTACTGA - GCTGTCCAGTGGAGGGCTCATGGGTGGGACATGGAAAAAGAGGCA 490
Qy 842 GCCTAGGCCCTGGGGAGCCAGTCCACTGAGCAAGCAAGGGACTGAGTGAAGCCCTTTTGCA 901
Db 491 GCCTAGGCCCTGGGGAGCCAGTCCACTGAGCAAGCAAGGGACTGAGTGAAGCCCTTTTGCA 550
Qy 902 CGAAAGGCTTAAGAAAAAGAAAAACCATTTCTAAACACACACAGAAACTGTCCAAATGCT 961
Db 551 GSAAGAGGCTTAAGAAAAAGAAAAACCATTTCTAAACACACACAGAAACTGTCCAAATGCT 610
Qy 962 TTGGGAACGTGTGTATTGCTTATAATGGGTCCCAAAATGGGTAACTTAGACTTCAGAG 1021
Db 611 TTGGGAACGTGTGTATTGCTTATAATGGGTCCCAAAATGGGTAACTTAGACTTCAGAG 670
Qy 1022 AGAATGAGCAGAGAGCAAGAGAAATCTGGCTGTCTCTTCCATTTTCATTTCTGTATCTC 1081
Db 671 AGAATGAGCAGAGAGCAAGAGAAATCTGGCTGTCTCTTCCATTTTCATTTCTGTATCTC 730
Qy 1082 AGGTGAGCTGTTAGAGGGAGACATTTAGAAAAAATGAACAA - CAAACAAATTAATAAT 1140
Db 731 AGGTGAACTGGTAAAGGGAGACATTTGAAAAAATGAACAAACCAAAACCAATTAATAAT 790
Qy 1141 GAGGTAC 1147
Db 791 GAGGTAC 797
```

RESULT 15

```
US-10-202-193-95
; Sequence 95, Application US/10202193
; Publication No. US20020192699A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jimmy
; APPLICANT: Astel, Jon H.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ford, Donna M.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; APPLICANT: Steinmann, Kathleen E.
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
; TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER
; FILE REFERENCE: PP-01532.103/200130.463D1
; CURRENT APPLICATION NUMBER: US/10/202,193
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1024)
; OTHER INFORMATION: n = A,T,C or G
```

US-10-202-193-95

```
Query Match 32.1%; Score 622; DB 13; Length 1024;
Best Local Similarity 95.0%; Pred. No. 8.1e-159;
Matches 690; Conservative 0; Mismatches 30; Indels 6; Gaps 5;

Qy 453 TTNIGTNTTTTTTTTTTTTCCGTCCTCCCAAGCTTTATCTGTCTTGACTTTTTTAAAAA 512
Db 101 TTTTTTTTTTTTTTTTTTTTCCGTCCTCCCAAGCTTTATCTGTCTTGACTTTTTTAAAAA 160
Qy 513 GTTTGGGGCAGATTCATGTAATGGGCTAAAGACATGCATTTTAAAACTAGGCAACTTC 572
Db 161 GTTTGGGGCAGATTCATGTAATGGGCTAAAGACATGCATTTTAAAACTAG--CAACTC 217
```

```
Qy 573 TTATTTCTTTCTTTAAATAATACATAGCATTAATCCCAATCCCTATTTAAAGACCTGAC 632
Db 218 TTATTTCTTTCTTTAAATAATACATAGCATTAATCCCAATCCCTATTTAAAGACCTGAC 277
Qy 633 AGCTTGAGAAGGTCACTACTGCAATTTATAGGACCTTCTGGTGGTTCGTCTGTACCTTTG 692
Db 278 AGCTTGAGAAGGTCACTACTGCAATTTATAGGACCTTCTGGTGGTTCGTCTGTACCTTTG 337
Qy 693 AAGTCTGCAAACTCTTGAGAACTTTTGCATGCAGAGGAGGTAAGAGGTAATTTGATTTTCA 752
Db 338 AAGTCTGCAAACTCTTGAGAACTTTTGCATGCAGAGGAGGTAAGAGGTAATTTGATTTTCA 397
Qy 753 CAGAGGAAGAAACACAGCGCAGAAATGAAGGGCCAGGCTTACTGAGGCTGTCCAGTGGAGGG 812
Db 398 CAGAGGAAGAAACACAGCGCAGAAATGAAGGGCCAGGCTTACTGA - GCTGTCCAGTGGAGGG 456
Qy 813 CTCATGGGTGGGACATGGAAAAAGAGGAGCCCTAGGCCCTGGGGAGCCAGTCCACTGAG 872
Db 457 CTCATGGGTGGGACATGGAAAAAGAGGAGCCCTAGGCCCTGGGGAGCCAGTCCACTGAG 516
Qy 873 CAAGCAAGGGACTGAGTGAAGCTTTTGCAGGAAAAAGGCTTAAGAAAAAGGAAACCAATTC 932
Db 517 CAAGCAAGGGACTGAGTGAAGCTTTTGCAGGAAAAAGGCTTAAGAAAAAGGAAACCAATTC 576
Qy 933 AAAAAACAACAAAGAAACTGTCCAAATGCTTTTGGGAACTGTGTTTATTTGCTTATAATGGGT 992
Db 577 AAAAAACAACAAAGAAACTGTCCAAATGCTTTTGGGAACTGTGTTTATTTGCTTATAATGGGT 636
Qy 993 CCCCCAAATGGGTAACTTAGACTTCAGAGAGAAATGAGCAGAGCAAGAGAGAAATCTGG 1052
Db 637 CCCCCAAATGGGTAACTTAGACTTCAGAGAGAAATGAGCAGAGCAAGAGAGAAATCTGG 696
Qy 1053 CTGTCTTCCATTTTCATTTCTGTTATCTCAGGTGAGCTGTAGAGGGGAGACATTAGAAA 1112
Db 697 CTG - CCTTCCATTTTCATTTCTGNTATCTCAGGTGAACTGGTANANGGGAGACATTNGAAA 755
Qy 1113 AAAATGAACAA - CAAAACAATTAATGAGGTAGCTGAGGCGCTGGGAGTCTCTTTGAC 1171
Db 756 AAAATGAACAAACCAAAACCAATTAATGAGGTACTTATGAGGTACCTTNGGNCCTGGGAAACACGCTTAAG 815
Qy 1172 TCCACT 1177
Db 816 GCGAAT 821
```

Search completed: August 26, 2003, 20:43:56  
Job time : 581 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2003, 13:36:36 ; Search time 4917 Seconds  
(without alignments)  
9569.548 Million cell updates/sec

Title: US-09-700-700-1  
Perfect score: 1936  
Sequence: 1 atggtatgccaaactaagt.....ggggcgccgcgactagtga 1936

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

- 1: em\_estba.\*
- 2: em\_esthum.\*
- 3: em\_estin.\*
- 4: em\_estnu.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_htc.\*
- 9: gb\_est1.\*
- 10: gb\_est2.\*
- 11: gb\_htc.\*
- 12: gb\_est3.\*
- 13: gb\_est4.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: em\_gss\_hum.\*
- 18: em\_gss\_inv.\*
- 19: em\_gss\_pin.\*
- 20: em\_gss\_vrt.\*
- 21: em\_gss\_fun.\*
- 22: em\_gss\_mam.\*
- 23: em\_gss\_mus.\*
- 24: em\_gss\_pro.\*
- 25: em\_gss\_rod.\*
- 26: em\_gss\_pbg.\*
- 27: em\_gss\_vrl.\*
- 28: gb\_gss1.\*
- 29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	728.4	37.6	742	13	BX102941
2	693.8	35.8	813	14	CB853124
3	689.6	35.6	851	13	BU601144
4	687.6	35.5	700	13	BU622196

C	5	643.2	33.2	674	12	BM983868
C	6	617	31.9	661	12	BQ020946
C	7	600	31.0	678	10	AW86523
C	8	599.2	31.0	608	13	BU619261
C	9	585.8	30.3	631	14	CA430903
C	10	582.2	30.1	599	9	AA813266
C	11	581.4	30.0	663	12	BQ018482
C	12	554.2	28.6	681	10	BE973555
C	13	550	28.4	563	9	AI989273
C	14	532	27.5	565	10	BF679599
C	15	521.8	27.0	538	9	AI620178
C	16	506.4	26.2	543	9	AA528104
C	17	492	25.4	504	9	AA531255
C	18	484.2	25.0	490	9	AA437224
C	19	482.2	24.9	1470	13	BQ230696
C	20	477	24.6	987	10	BF965018
C	21	448	23.1	496	9	AA516531
C	22	445	23.0	456	9	AA442287
C	23	443	22.9	472	9	AA640241
C	24	438.8	22.7	455	9	AA972883
C	25	436	22.5	488	10	BE772782
C	26	419.4	21.7	434	9	AI989281
C	27	416.8	21.5	435	9	AA809587
C	28	416	21.5	416	9	AA707108
C	29	414	21.4	414	9	AI139031
C	30	413.4	21.4	415	9	AI052805
C	31	413.4	21.4	418	9	AI791607
C	32	409	21.1	409	9	AI571633
C	33	407.4	21.0	423	9	AI685682
C	34	405.2	20.9	630	9	AI557413
C	35	404.2	20.9	445	10	BF855203
C	36	394.8	20.4	419	9	AA229495
C	37	389.2	20.1	607	12	BI855676
C	38	386	19.9	411	9	AA640928
C	39	385.4	19.9	660	9	AI525269
C	40	384.2	19.8	423	9	AA531606
C	41	380.8	19.7	403	9	AA493522
C	42	379	19.6	445	9	AA533772
C	43	375.2	19.4	431	10	BF855210
C	44	374.6	19.3	409	10	BF855199
C	45	368.8	19.0	386	9	AI027196

ALIGNMENTS

RESULT 1  
BX102941  
LOCUS  
DEFINITION BX102941 Soares testis NHT Homo sapiens cdna clone IMAGp998M201862  
ACCESSION BX102941  
VERSION BX102941  
KEYWORDS EST:  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 742)  
AUTHORS Ebert, L., Helli, O., Hennig, S., Neubert, P., Partsch, E., Peters, M., Radelof, U., Schneider, D. and Korn, B.  
TITLE Human Unigeneset - RZPD3 (RZPDLIB No. 972)  
JOURNAL Unpublished  
COMMENT Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
RZPD; IMAGp998M201862.  
RZPDLIB; I.M.A.G.E. cdna Clone Collection;  
Human Unigeneset - RZPD3 (RZPDLIB No. 972)  
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany

Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
[www.rzpd.de](http://www.rzpd.de)

This clone is available royalty-free from RZPD;

contact RZPD (clone@rzd.de) for further information. Seq primer: M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.

FEATURES	SOURCE
1. The first two columns are labeled "FEATURES" and "SOURCE".	
2. The third column is labeled "FEATURES".	
3. The fourth column is labeled "SOURCE".	
4. The fifth column is labeled "FEATURES".	
5. The sixth column is labeled "SOURCE".	
6. The seventh column is labeled "FEATURES".	
7. The eighth column is labeled "SOURCE".	
8. The ninth column is labeled "FEATURES".	
9. The tenth column is labeled "SOURCE".	
10. The eleventh column is labeled "FEATURES".	
11. The twelfth column is labeled "SOURCE".	
12. The thirteenth column is labeled "FEATURES".	
13. The fourteenth column is labeled "SOURCE".	
14. The fifteenth column is labeled "FEATURES".	
15. The sixteenth column is labeled "SOURCE".	
16. The seventeenth column is labeled "FEATURES".	
17. The eighteenth column is labeled "SOURCE".	
18. The nineteenth column is labeled "FEATURES".	
19. The twentieth column is labeled "SOURCE".	
20. The twenty-first column is labeled "FEATURES".	
21. The twenty-second column is labeled "SOURCE".	
22. The twenty-third column is labeled "FEATURES".	
23. The twenty-fourth column is labeled "SOURCE".	
24. The twenty-fifth column is labeled "FEATURES".	
25. The twenty-sixth column is labeled "SOURCE".	
26. The twenty-seventh column is labeled "FEATURES".	
27. The twenty-eighth column is labeled "SOURCE".	
28. The twenty-ninth column is labeled "FEATURES".	
29. The thirtieth column is labeled "SOURCE".	
30. The thirty-first column is labeled "FEATURES".	
31. The thirty-second column is labeled "SOURCE".	
32. The thirty-third column is labeled "FEATURES".	
33. The thirty-fourth column is labeled "SOURCE".	
34. The thirty-fifth column is labeled "FEATURES".	
35. The thirty-sixth column is labeled "SOURCE".	
36. The thirty-seventh column is labeled "FEATURES".	
37. The thirty-eighth column is labeled "SOURCE".	
38. The thirty-ninth column is labeled "FEATURES".	
39. The fortieth column is labeled "SOURCE".	
40. The forty-first column is labeled "FEATURES".	
41. The forty-second column is labeled "SOURCE".	
42. The forty-third column is labeled "FEATURES".	
43. The forty-fourth column is labeled "SOURCE".	
44. The forty-fifth column is labeled "FEATURES".	
45. The forty-sixth column is labeled "SOURCE".	
46. The forty-seventh column is labeled "FEATURES".	
47. The forty-eighth column is labeled "SOURCE".	
48. The forty-ninth column is labeled "FEATURES".	
49. The fiftieth column is labeled "SOURCE".	
50. The fifty-first column is labeled "FEATURES".	
51. The fifty-second column is labeled "SOURCE".	
52. The fifty-third column is labeled "FEATURES".	
53. The fifty-fourth column is labeled "SOURCE".	
54. The fifty-fifth column is labeled "FEATURES".	
55. The fifty-sixth column is labeled "SOURCE".	
56. The fifty-seventh column is labeled "FEATURES".	
57. The fifty-eighth column is labeled "SOURCE".	
58. The fifty-ninth column is labeled "FEATURES".	
59. The sixtieth column is labeled "SOURCE".	
60. The sixty-first column is labeled "FEATURES".	
61. The sixty-second column is labeled "SOURCE".	
62. The sixty-third column is labeled "FEATURES".	
63. The sixty-fourth column is labeled "SOURCE".	
64. The sixty-fifth column is labeled "FEATURES".	
65. The sixty-sixth column is labeled "SOURCE".	
66. The sixty-seventh column is labeled "FEATURES".	
67. The sixty-eighth column is labeled "SOURCE".	
68. The sixty-ninth column is labeled "FEATURES".	
69. The seventieth column is labeled "SOURCE".	
70. The seventy-first column is labeled "FEATURES".	
71. The seventy-second column is labeled "SOURCE".	
72. The seventy-third column is labeled "FEATURES".	
73. The seventy-fourth column is labeled "SOURCE".	
74. The seventy-fifth column is labeled "FEATURES".	
75. The seventy-sixth column is labeled "SOURCE".	
76. The seventy-seventh column is labeled "FEATURES".	
77. The seventy-eighth column is labeled "SOURCE".	
78. The seventy-ninth column is labeled "FEATURES".	
79. The eightieth column is labeled "SOURCE".	
80. The eighty-first column is labeled "FEATURES".	
81. The eighty-second column is labeled "SOURCE".	
82. The eighty-third column is labeled "FEATURES".	
83. The eighty-fourth column is labeled "SOURCE".	
84. The eighty-fifth column is labeled "FEATURES".	
85. The eighty-sixth column is labeled "SOURCE".	
86. The eighty-seventh column is labeled "FEATURES".	
87. The eighty-eighth column is labeled "SOURCE".	
88. The eighty-ninth column is labeled "FEATURES".	
89. The ninetieth column is labeled "SOURCE".	
90. The ninety-first column is labeled "FEATURES".	
91. The ninety-second column is labeled "SOURCE".	
92. The ninety-third column is labeled "FEATURES".	
93. The ninety-fourth column is labeled "SOURCE".	
94. The ninety-fifth column is labeled "FEATURES".	
95. The ninety-sixth column is labeled "SOURCE".	
96. The ninety-seventh column is labeled "FEATURES".	
97. The ninety-eighth column is labeled "SOURCE".	
98. The ninety-ninth column is labeled "FEATURES".	
99. The hundredth column is labeled "SOURCE".	

```

1. /42
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE998M201862 ; IMAGE:757435"
/sex="male"
/lab_host="DH10B"

```

/clone lib="Soares testis\_NHT"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from Clontech Laboratories  
 , Inc., and primed with a Not I - oligo(dT) primer [5',  
 TGTACCACTGAAGTGGAGCGCGCCCAATTTTCTTTT 3'].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization to Cot5, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."  
 224 a 144 c 185 g 188 t 1 others

Query Match 37.6%; Score 728.4; DB 13;  
Best Local Similarity 99.6%; Pred. No. 1.9e-106;  
Matches 740. Cases

Matches	740;	Conservative	0;	Mismatches	2;	Indels	1;	Gaps	1;
---------	------	--------------	----	------------	----	--------	----	------	----

QY  
601 ATTAAATCCCAAATCCTATTTTAAGACCTGACCGCTTCACACACCCTCAC

660

Db  
1 ATTAAATCCCAATCCTATTTAAGACCTGACAGCTTGAGAAAGGTCACTACTGCATTAT 60

QY 661 AGGACCTTCTGGTGCTTCATCATCAGCCAGCGA

5' AGACCTTCCTGGGTCGCTGTACGTTGAAGTCTGACAAATCCTTGAGAAATCTTTGC 720

D<sub>b</sub>

61 AGGACCTTCTGGTGGTTCTGCTGTACGTTTGAAGTCTGACAAATCCTTGAGAAATCTTTGC 120

721 ATCC 35061 120

QY  
/21 ATGCAGAGGAGGTAAGAGGTATTGGATTTCACAGAGGAACAACAGCGCAGAAATGAAG 780

**D<sub>b</sub>**      121 ATGCAGAGGAGTAAAGAGGTAATTGCA TTTTCACAGAGCGACACACACCCCTATCCTA

Q: CAGCGGAAACACAGCGAGAATGAAG 180

781 GCCAGGCTTACTGAGGCTGTCCAGTGGAGGGCTCATGGTGGGACATGGAAGAAGGC 840

Db  
181 GGCCAGGCTTACTGA - GGTCTCCACCTCCGCCCTATTCCTCCT

101 GGCAGGCTTACTGAGTGGTGGACATGGAAAGAGGC 239

841 AGCCTAGGCCCTGGGGAGCCAGTCCACTGAGCAAGCAAGGACTGAGTGAGCCTTTTGC 900

db  
240

DD 240 AGCCCTAGGCCCTGGGGAGCCCAGTCCACTGAGCAAGCAAGGACTGAGTGAGCCTTTTGC 299

QY  
901 AGGAAAGGCTAAGAAAAAGGAAAACCATTCATAAACCAACAACAATCCTCCATTGG

|||||TCTGTTTCATCCCATCTCATTACACAAAGAAAC"GTCCAAATGC 960

300 AGGAAAAGGCTAGAAAAGGAAAACCATTTCTAAAACACAAGAACTGTCCAAATGC 359

QY 961 TTGTGGGAAC TGTGTTTATATTCCCTATA TACCGTCAGCCT

201 11GGGAC1GG111A11GGCTAA1AATGGGTCCCCAAAATGGGTAACCTAGACTTCAGA 1020

Db  
360 TTGGGAACGTGTTTATTGCCCTATAATGGGTCCCCAAATGGGTAACTTAGACTTCAGA 419

Qv 1021 CACCACTCCGCGCAGGGTGTTCATTTTGGTAACCTAGACTTCAGA 419

1021 GAGAA'GAGCAGAGCAAGGAGAAATCTGGCTGTCCTCCATTTTCATTCTGTATCT 1080

D**b**

420 GAGAA TGAGC AGAGCA ACGAGA ATCTGG CTGTC CATTCC ATTATC ATTCTT CA

Q: 1001 CAGAGGCTTCCGACGCTTCCGCTGCTATCT 479

1081 CAGGTGAGCTGGTAGAGGGGAGACATTAGAAAAAAATGAAACAACAAAACAATTACTAAT 1140

db  
480 CAGGTGAGCTGGTAGAGGGGACGCATTACAAATAATTCCTCCTCC

539

1141 GAGTACGCTGAGGCCCTGGAGTCTCTTGACTCCACTACTTAATCCGTTTAGTGAGAAA 1200

db

540 GAGCTACCCCTCACCCTCCGCTATGGTATT  
|||||  
|||||

1200 CCGTTTAGTGAGAAH

340 GAGTACGC TGAGGCCCTGGGAGTCTCTTGACTCCACTACTTAATTCGTTTAGTGAGAA 599

Qy 1201 CCTTCAATTTCTTTATTAGAAGGCCAGCTTACTGTGTGGCAAAATTCCCCACCT 1200

-----GCGAGCCATACCTGTGGTGGCAAAATGGCCAACAT 1260





```

REFERENCE
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        1 (bases 1 to 674)
JOURNAL      Bonaldo,M.F., Lennon,G. and Soares,M.B.
MEDLINE      Normalization and subtraction: two approaches to facilitate gene
PUBMED       discovery
COMMENT      97044477
              Genome Res. 6 (9), 791-806 (1996)
              8889548
              Contact: McCray, PB
              McCray Lab
              University of Iowa
              2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
              Tel: 319 356 4866
              Fax: 319 356 7171
              Email: paul-mccray@iowa.edu
              Tissue Procurement: Dr. M. J. Welsh, University of Iowa
              cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
              cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: Researchers may obtain clones from Research
              Genetics (www.resgen.com) or from Open Biosystems
              (www.openbiosystems.com).
              The following repetitive elements were found in this cDNA
              sequence: 1-21, >AT rich#Low_complexity (matched complement)
              Seq primer: M13 FORWARD
              POLYA=Yes.

FEATURES
source
1. .674
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="UI-CF-DUI-aaw-b-04-0-UI"
   /tissue_type="Primary Lung Epithelial Cells"
   /dev_stage="Adult"
   /lab_host="PH10B (Life Technologies) (T1 phage resistant)"
   /clone_lib="UI-CF-DUI"
   /notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
   modified polylinker; Site 1: EcoR I; Site 2: Not I;
   UI-CF-DUI is a normalized cDNA library containing the
   following tissue(s): Primary Lung Epithelial Cells The
   library was constructed according to Bonaldo, Lennon and
   Soares, Genome Research, 6:791-806, 1996. First strand
   cDNA synthesis was primed with an oligo-dT primer
   containing a Not I site. Double stranded cDNA was ligated
   to an EcoR I adaptor, digested with Not I, and cloned
   directionally into pT7T3-Pac vector. The oligonucleotide
   used to prime the synthesis of first-strand cDNA contains
   a library tag sequence that is located between the Not I
   site and the (dT)18 tail. The sequence tag for this
   library is GGCTGTAGGC.
   TAG_LTB=UI-CF-DUI
   TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
   TAG_SEQ=GGCTGTAGGC"

BASE COUNT      221 a 111 c 144 g 198 t
ORIGIN

Query Match      33.2%; Score 643.2; DB 12; Length 674;
Best Local Similarity 98.8%; Pred. No. 7.7e-93;
Matches 669; Conservative 0; Mismatches 4; Indels 4; Gaps 2;

QY 1241 GGTGGCAAAATGCCAACATAGATTAAAGAGTTGGCAATTCACCCCAATTTCTGT 1300
DB 674 GGTGGCAAAATGCCAACATAGATTAAAGAGTTGGCAATTTTCCACCCCAATTTCTGT 615
QY 1301 GGTGGGCTCCACATGCAATGTTCAATGCCACGCTGCTGACACCGCGAGTACT 1360
DB 614 GGTGGGCTCCACATGCAATGTTCAATGCCACGCTGCTGACACCGCGAGTACT 555
QY 1361 AGCCAGCAAAAGGAGGAGTGAATTCGTTCTGCTCTTTACATTTCTTTTAAAA 1420
DB 554 AGCCAGCAAAAGGAGGAGTGAATTCGTTCTGCTCTTTACATTTCTTTTAAAA 495
QY 1421 TAAGCATTTAGTGCTCAGTCCCTACTGTAGTACTCTTTCTCTCCCTCTCTGAAATTTAAT 1480

```

```

Db 494 TAAGCATTTAGTGCTCAGTCCCTACTGTAGTACTCTTTCTCTCCCTCTCTGAAATTTAAT 435
QY 1481 TCTTTCAACTTGCATTTGCAAGGATTACACATTTTCACTGTGATGTATATTGTGTGCAG 1540
Db 434 TCTTTCAACTTGCATTTGCAAGGATTACACATTTTCACTGTGATGTATATTGTGTGCA- 376
QY 1541 NGAAAAGAAAAAGTGTCTTTGTTTAAATTAATCTTGGTTTGAATFCCATCTTGTCTTTT 1600
Db 375 --AAAAAAGAAAAAGTGTCTTTGTTTAAATTAATCTTGGTTTGAATFCCATCTTGTCTTTT 318
QY 1601 CCCCAATTCGNACTAGTCAATTAACCCATCTCTGAACCTGGTAGAAAAACATCTCAAGAGCTA 1660
Db 317 CCCCAATTCGNACTAGTCAATTAACCCATCTCTGAACCTGGTAGAAAAACATCTCAAGAGCTA 258
QY 1661 GTCATATCAGCATCTCAGAGGTGAATTTGGATGTTCTCAGAACCATTTTCAACCAGACAGCC 1720
Db 257 GTCATATCAGCATCTCAGAGGTGAATTTGGATGTTCTCAGAACCATTTTCAACCAGACAGCC 198
QY 1721 TGTTCCTATCCTGTTTAAATTAATTAATTTGGGTTTCTTACATGCATTAACAAACCCCTGCTC 1780
Db 197 TGTTCCTATCCTGTTTAAATTAATTAATTTGGGTTTCTTACATGCATTAACAAACCCCTGCTC 138
QY 1781 CAATCTGCACATAAAAGTCTGTGACTTGAAGTTTAGTTCAGCAGCACCCCA-CCAAACTTTA 1839
Db 137 CAATCTGCACATAAAAGTCTGTGACTTGAAGTTTAGTTCAGCAGCACCCCACTTTA 78
QY 1840 TTTTCTATGTGTTTGTGCAACATATGATGTTTGTGAAATTAAGTACCCATGCTTTTA 1899
Db 77 TTTTCTATGTGTTTGTGCAACATATGATGTTTGTGAAATTAAGTACCCATGCTTTTA 18
QY 1900 TTAATAAANAATAAATAA 1916
Db 17 AAAAAAAAAAAAAAAAAA 1

RESULT 6
BQ020946/c
LOCUS
DEFINITION
IMAGE:5828837 3', mRNA sequence.
ACCESSION
BQ020946
VERSION
BQ020946.1 GI:19756232
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 661)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
The following repetitive elements were found in this cDNA
sequence: 1-30, >AT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1. .661
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:5828837"
   /tissue_type="Metastatic Chondrosarcoma"
   /dev_stage="Adult"
FEATURES
source

```

```

/lab host="DH10B (Life Technologies)"
/clone lib="NCI CGAP DH1"
/notes="Organ: Lung; Vector: pTT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP DH1 is a normalized cDNA library containing the
following tissue(s): VS-8 Cell line from Metastatic
Chondrosarcoma in Lung. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dr primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pTT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is AGATCATGTC.
TAG LIB-UI-H-DH1
TAG_TISSUE=Lung
TAG_SEQ=AGATCATGTC"
BASE COUNT      220 a 106 c 139 g 194 t 2 others
ORIGIN
Query Match      31.9%; Score 617; DB 12; Length 661;
Best Local Similarity 98.2%; Pred. No. 1.2e-88;
Matches 644; Conservative 0; Mismatches 8; Indels 4; Gaps 2;
QY 1261 AAGTTAATAGAAAGTTGGCCAAATTCACCCCAATTTCTGTGGTTGGGCTCCCATTTGCA 1320
DB 661 AAGTTAATAGAAAGTTGGCCAAATTCACCCCAATTTCTGT- GTTGGGCTCCCATTTGCA 603
QY 1321 ATGTTCAATGCCAGTGTCTGTACACGACCGAGTACTAGCCAGCACAAAAGGAGGG 1380
DB 602 ATGTTCAATGCCAGTGTCTGTACACGACCGAGTACTAGCCAGCACAAAAGGAGGG 543
QY 1381 TAGCCTGAATTCCTTCTGCTCTTTACATTTCTTTTAAATAAGCATTTAGTCTCAGTC 1440
DB 542 TAGCCTGAATTCCTTCTGCTCTTTACATTTCTTTTAAATAAGCATTTAGTCTCAGTC 483
QY 1441 CCTACTGAGTACTCTTCTCTCCCTCTCTGAAATTAATTTCTTCAACTTGCATTTGC 1500
DB 482 CCTACTGAGTACTCTTCTCTCCCTCTCTGAAATTAATTTCTTCAACTTGCATTTGC 423
QY 1501 AAGGATTACATTTTCACTGTGATGATATATGTTGTCGAGNAGAAAGAAAGTGTCTT 1560
DB 422 AAGGATTACATTTTCACTGTGATGATATATGTTGTC- - - - -AAAAAAGAAAGTGTCTT 366
QY 1561 TGTTTAAATTAATTTGTTTGTGAATTCATCTTGTCTTTTCCCATTTGGAATAGTCAAT 1620
DB 365 TGTTTAAATTAATTTGTTTGTGAATTCATCTTGTCTTTTCCCATTTGGAATAGTCAAT 306
QY 1621 AACCCATCTCTGAAGTGTAGAAAACATCTGAAGAGTACTATCAGCATCTGACAGG 1680
DB 305 AACCCATCTCTGAAGTGTAGAAAACATCTGAAGAGTACTATCAGCATCTGACAGG 246
QY 1681 TGAATTTGATGTTCTCAGAACCAATTTACCCAGACAGCTGTTTCTATCTCTGTTTAATA 1740
DB 245 TGAATTTGATGTTCTCAGAACCAATTTACCCAGACAGCTGTTTCTATCTCTGTTTAATA 186
QY 1741 AATTAGTTGGGTTCTCTACATGCAACAAACCCCTGCTCCAATCTGTGCATATAAAGTC 1800
DB 185 AATTAGTTGGGTTCTCTACATGCAACAAACCCCTGCTCCAATCTGTGCATATAAAGTC 126
QY 1801 TGTGACTTGAAGTTTGTAGTCAGACCCCAACCACTTATTTTCTATGTTGTTTTTGA 1860
DB 125 TGTGACTTGAAGTTTGTAGTCAGACCCCAACCACTTATTTTCTATGTTGTTTTTGA 66
QY 1861 ACATATGAGTTTGTAAATAAAGTACCCATGTCCTTTTATTAATAAATAAATAA 1916
DB 65 ACATATGAGTTTGTAAATAAAGTACCCATGTCCTTTTATTAATAAATAAATAA 10

LOCUS      AW886523      678 bp      mRNA      linear      EST 23-MAY-2000
DEFINITION RCI-OT0083-220300-021-f10 OT0083 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW886523
VERSION    AW886523.1 GI:8048535
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 678)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baig,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
PUBMED     10737800
COMMENT    Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel.: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2=RC1-OT0083-220
300-021-f10&t3=2000-03-22&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 649.
Location/Qualifiers
1. 678
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="OT0083"
/notes="Organ: ovary; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      196 a 145 c 163 g 174 t
ORIGIN
Query Match      31.0%; Score 600; DB 10; Length 678;
Best Local Similarity 97.4%; Pred. No. 5.9e-86;
Matches 631; Conservative 0; Mismatches 15; Indels 2; Gaps 2;
QY 781 GGCAGAGCTTACTGAGGCTGTCCAGTGGAGGGCTCATGGTGGACATGGAAGAAGGC 840
DB 2 GGCAGAGCTTACTGA-GCTGTCCAGTGGAGGGCTCATGGTGGACATGGAAGAAGGC 60
QY 841 AGCCTAGGCCCTGGGGAGCCAGTCCACTGAGCAAGGAGCTGAGTGGAGCCTTTTGC 900
DB 61 AGCCTAGGCCCTGGGGAGCCAGTCCACTGAGCAAGGAGCTGAGTGGAGCCTTTTGC 120
QY 901 AGGAAAGGCTTAAAGAAAGGAAACCATTTCTAAACACAAAGAACTGTCCCAATGC 960
DB 121 AGGAAAGGCTTAAAGAAAGGAAACCATTTCTAAACACAAAGAACTGTCCCAATGC 180
QY 961 TTTGGGAACCTGTGTTTATTGCTTATATGGTGTCCCAAAATGGGTAACTTAGACTTCAGA 1020
DB 181 TTTGGGAACCTGTGTTTATTGCTTATATGGTGTCCCAAAATGGGTAACTTAGACTTCAGA 240
QY 1021 GAGAATGAGCAGAGACAAAGAGAAATCTGGGCTGCTTCCATTTTCACTTGTATCT 1080

```

Db 241 GAGAAATGAGCAGAGACAAAGGAGAAATCTGGCTGCTCTTCCATTTTCAATTTCTGTTATCT 300  
QY 1081 CAGGTGAGCTGGTAGAGGGGAGACATTAGAAAAAATGAAACAAACAAACAAATTAAT 1140  
Db 301 CAGGTGAGCTGGTAGAGGGGAGACATTAGAAAAAATGAAACAAACAAACAAATTAAT 360  
QY 1141 GAGGTGAGCTGGAGGCTGGAGGCTCTTGACTCCACTACTTAATTCGTTTGTAGTGA 1200  
Db 361 GAGGTGAGCTGGAGGCTGGAGGCTCTTGACTCCACTACTTAATTCGTTTGTAGTGA 420  
QY 1201 CTTTCAATTTCTTTTATAGAGGGCAGCTACTGTTGGTGGCAAAATTCGCAACAT 1260  
Db 421 CTTTCAATTTCTTTTATAGAGGGCAGCTACTGTTGGTGGCAAAATTCGCAACAT 480  
QY 1261 AAGTTAATAGAAATGGGCAATTTTCAACCCCATTTTCTGGTGGTTCGCTCCATTTGCA 1320  
Db 481 AAGTTAATAGAAATGGGCAATTTTCAACCCCATTTTCTGGTGGTTCGCTCCATTTGCA 540  
QY 1321 ATGTTCAATGCCAGCTGCTGCTGACACCGACCGAGTACTAGCCAGCACAAAGGCAGG 1380  
Db 541 ATGTTCAATGCCAGCTGCTGCTGACACCGACCGAGTACTAGCCAGCACAAAGGCAGG 600  
QY 1381 TAGCCTGAATGCTTTCTGCTCTTTTACATTTCTTTTAAATAGCAAT 1428  
Db 601 TAGCCTGAATGCTGCTGCTCTTTTCAATTTCTTTTAAATAGCAAT 647

RESULT 8  
BU619261/c  
LOCUS  
DEFINITION  
UI-H-FH1-bfr-b-10-0-UI.s1 NCI CGAP FH1 Homo sapiens cDNA clone  
UI-H-FH1-bfr-b-10-0-UI 3', mRNA sequence.  
ACCESSION  
BU619261  
VERSION  
BU619261.1 GI:23285476  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 608)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
AUTHORS  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL  
Unpublished  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)  
Tissue Procurement: James Martin  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
Seq primer: M13 FORWARD  
POLYA=Yes.

FEATURES  
source  
1..608  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-FH1-bfr-b-10-0-UI"  
/issue\_type="Cell Line"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI CGAP\_FH1"  
/notes="Organ: Chondrosarcoma; Vector: pT73-Pac (Pharmacia)  
I; with a modified polylinker; Site 1: EcoR I; Site 2: Not  
I; NCI CGAP FH1 is a normalized cDNA library obtained from  
a cell line derived from grade I chondrosarcoma tissue.  
The library was constructed and normalized according to  
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,  
1996. First strand cDNA synthesis was primed with an  
oligo-dT primer containing a Not I site. Double stranded

cDNA was ligated to an EcoR I adaptor, digested with Not  
I, and cloned directionally into pT73-Pac vector. The  
oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (GT)<sub>18</sub> tail. The  
sequence tag for this library is AGAATCCGCG. The cell line  
was provided by Dr. James Martin from the University of  
Iowa.  
TAG LIB=UI-H-FH1  
TAG TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1  
Chondrosarcoma  
TAG\_SEQ=AGAATCCGCG"  
BASE COUNT 184 a 121 c 124 g 179 t  
ORIGIN  
Query Match 31.0%; Score 599.2; DB 13; Length 608;  
Best Local Similarity 99.3%; Pred. No. 8.3e-86;  
Matches 601; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 949 CTGTCCAAATGCTTTGGAACTGTGTTTATTCCTATATATGGGTCCCAAAATGGGTAAAC 1008  
Db 608 CTGTCCAAATGCTTTGGAACTGTGTTTATTCCTATATATGGGTCCCAAAATGGGTAAAC 549  
QY 1009 CTAGACTTCAGAGAAATGAGCAGAGAGCAAGGAGAAATCTGGCTGTCTTCCATTTTC 1068  
Db 548 CTAGACTTCAGAGAAATGAGCAGAGAGCAAGGAGAAATCTGGCTGTCTTCCATTTTC 489  
QY 1069 ATTCTGTATCTCAGTGTGAGTGGTAGAGGGGAGACATTAGAAAAAATGAAACACAA 1128  
Db 488 ATTCTGTATCTCAGTGTGAGTGGTAGAGGGGAGACATTAGAAAAAATGAAACACAA 429  
QY 1129 ACAATTAATAGAGTACGCTGAGCGCTGGGAGTCTCTTGACTCCACTACTTAATTCGG 1188  
Db 428 ACAATTAATAGAGTACGCTGAGCGCTGGGAGTCTCTTGACTCCACTACTTAATTCGG 369  
QY 1189 TTTAGTGAGAAACCTTTCAATTTTCTTTTATAGAGGGCCAGCTTACTGTTGGTGGCAA 1248  
Db 368 TTTAGTGAGAAACCTTTCAATTTTCTTTTATAGAGGGCCAGCTTACTGTTGGTGGCAA 309  
QY 1249 AATTCGCCAACATAAGTAAATAGAAAGTGGCCAAATTTTCAACCCCATTTCTGTTGGTGG 1308  
Db 308 AATTCGCCAACATAAGTAAATAGAAAGTGGCCAAATTTTCAACCCCATTTCTGTTGGTGG 249  
QY 1309 CTCCACATTTGCAATGTTTCAATGCCAGTCTGCTGACACCGACCGAGTACTAGCCAGCA 1368  
Db 248 CTCCACATTTGCAATGTTTCAATGCCAGTCTGCTGACACCGACCGAGTACTAGCCAGCA 189  
QY 1369 CAAAAGGCGAGGTAGCCTGAATGCTTCTGCTCTTTTACATTTCTTTTAAATAAGCAAT 1428  
Db 188 CAAAAGGCGAGGTAGCCTGAATGCTTCTGCTCTTTTACATTTCTTTTAAATAAGCAAT 129  
QY 1429 TAGTGCTCAGTCCCTACTGAGTACTCTTCTCTCCCTCCCTCTGAATTAATTTCTTCAA 1488  
Db 128 TAGTGCTCAGTCCCTACTGAGTACTCTTCTCTCCCTCCCTCTGAATTAATTTCTTCAA 69  
QY 1489 CTTGCAATTTGCAAGGATTACACATTTCACTGTGATGATATTTGTTGCGAGGAAAGCA 1548  
Db 68 CTTGCAATTTGCAAGGATTACACATTTCACTGTGATGATATTTGTTGCGAGGAAAGCA 9  
QY 1549 AAAAA 1553  
Db 8 AAAAA 4  
RESULT 9  
CA430903/c  
LOCUS  
DEFINITION  
UI-H-FH1-bge-i-17-0-UI.s1 NCI CGAP FH1 Homo sapiens cDNA clone  
UI-H-FH1-bge-i-17-0-UI 3', mRNA sequence.  
ACCESSION  
CA430903  
VERSION  
CA430903.1 GI:24793629  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 631)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: James Martin  
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, bento-soares@uiowa.edu  
Seq primer: M13 FORWARD  
POLYA=Yes.

## JOURNAL

## COMMENT

## FEATURES

source

Location/Qualifiers  
1..631

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="UI-H-FL1-bge-i-17-0-UI"

/tissue\_type="Cell lines"

/dev\_stage="Adult"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="NCI CGAP FL1"

/note="Organ: Chondrosarcoma; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FL1 is a normalized cDNA library derived from a pool of mRNA obtained from 4 cell lines from grade III chondrosarcoma tissues. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GAGTGGTG. The cell lines were provided by Dr. James Martin from the University of Iowa.  
TAG LIB=UI-H-FL1  
TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix  
TAG\_SEQ=GAGTGGTG"

## BASE COUNT

ORIGIN 210 a 102 c 136 g 182 t 1 others

## Query Match

Best Local Similarity 30.3%; Score 585.8; DB 14; Length 631;

Matches 622; Conservative 0; Mismatches 3; Indels 5; Gaps 3;

QY 1288 CCCCATTTCTGTGGTTGGCTCCCATGTCGATGTTCAATGCCACGTCGTGACAC 1347

DB 631 CCCCATTTCTGTGGTTGGCTCCCATGTCGATGTTCAATGCCACGTCGTGACAC 572

QY 1348 CGACCGGAGTACTAGCCAGCACAAAAG-GCAGGCTAGCTGAAATGCTTCTGCTCTTTA 1406

DB 571 CGACCGGAGTACTAGCCAGCACAAAAGGCGAGGTAGCTGAAATGCTTCTGCTCTTTA 512

QY 1407 CATTCTTTTAAATAGCATTTAGTGTCTAGTCCCTACTGAGTACTCTTCTCCCT 1466

DB 511 CATTCTTTTAAATAGCATTTAGTGTCTAGTCCCTACTGAGTACTCTTCTCCCT 452

QY 1467 CTTCTGAATTTAATCTTTCACTTTCGAATTTGCAAGGATTTACATTTCACTGTGATG 1526

DB 451 CTTCTGAA-TTAATCTTTCACTTTCGAATTTGCAAGGATTTACATTTCACTGTGATG 393

QY 1527 ATATTGTTGTCAGNAGAAAGAAAGTGTCTTTGTTTAAATTAATCTGTTTGAAT 1586

DB 392 ATATTGTTGTC---AAAAAAAAGTGTCTTTGTTTAAATTAATCTGTTTGAAT 336

QY

1587 CCATCTTGCTTTTCCCATTGGAACTAGTCAATTAAACCATCTCTGAACTGGTAGAAAA 1646

DB

335 CCATCTTGCTTTTCCCATTGGAACTAGTCAATTAAACCATCTCTGAACTGGTAGAAAA 276

QY

1647 CATCTGAAGAGCTAGTCTATCAGCATCTCAGAGTGAATGGATGTTCTCAGAACCAT 1706

DB

275 CATCTGAAGAGCTAGTCTATCAGCATCTCAGAGTGAATGGATGTTCTCAGAACCAT 216

QY

1707 TCACCAGACAGCTGTTTCTATCCCTGTTTATAAATAGTTTGGGTTCTCTACATGCAT 1766

DB

215 TCACCAGACAGCTGTTTCTATCCCTGTTTATAAATAGTTTGGGTTCTCTACATGCAT 156

QY

1767 AACAAACCTGCTCCAACTCTGCATCAATAAAGTCTGTGACTTGAAGTTTGTGACACCC 1826

DB

155 AACAAACCTGCTCCAACTCTGCATCAATAAAGTCTGTGACTTGAAGTTTGTGACACCC 96

QY

1827 CCACCAACTTTATTTTCTATGTTGTTTTCGCAACATATGAGTGTGTTTGAATAAAGT 1886

DB

95 CCACCAACTTTATTTTCTATGTTGTTTTCGCAACATATGAGTGTGTTTGAATAAAGT 36

QY

1887 ACCCATGCTTTTATTAATAAATAAATAAATAA 1916

DB

35 ACCCATGCTTTTATTAATAAATAAATAAATAA 6

## RESULT 10

AA813266/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AA813266 599 bp mRNA linear EST 31-DEC-1998  
aj44903.sl Soares\_testis\_NHT Homo sapiens cDNA clone 1393204 3',  
mRNA sequence.

AA813266  
AA813266.1 GI:2883251  
EST.

Homo sapiens (human)  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 599)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished

Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo  
, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Insert length: 982 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 503.

Location/Qualifiers  
1..599

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="1393204"

/sex="male"

/lab\_host="DH10B"

/clone\_lib="Soares testis\_NHT"

/note="Vector: pT73-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was prepared from mRNA obtained from Clontech Laboratories  
, Inc. and primed with a Not I - oligo(dT) primer [5',  
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization to Cot5, and was

## FEATURES

source

```

BASE COUNT      202 a      98 c      131 g      168 t
ORIGIN
Query Match      30.1%; Score 582.2; DB 9; Length 599;
Best Local Similarity 99.2%; Pred.No. 4.3e-83;
Matches 596; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 1307 GGCTCCACATTTGCAATGTTCAATGCCAGCTGCTGCTGACACCGGAGTACTAGCCAG 1366
Db 599 GGCTCCACATTTGCAATGTTCAATGGCCAGCTGCTGCTGACACCGGAGTACTAGCCAG 540

QY 1367 CACAAAAGGAGGAGTACGCTGAATGCTTCTGCTCTTTACATTTCTTTAAATAAGCA 1426
Db 539 CACAAAAGGAGGAGTACGCTGAATGCTTCTGCTCTTTACATTTCTTTAAATAAGCA 480

QY 1427 TTTAGTGCTCAGTCCCTACTGAGTACTCTTTCTCTCCCTCCCTCCTCAATTTAATCTTTC 1486
Db 479 TTTAGTGCTCAGTCCCTACTGAGTACTCTTTCTCTCCCTCCCTCCTCAATTTAATCTTTC 420

QY 1487 AACTTGCAATTTGCAAGGATTACACATTTCACTGTGATGTATATTTGTCAGNGAAAA 1546
Db 419 AACTTGCAATTTGCAAGGATTACACATTTCACTGTGATGTATATTTGTCAGNGAAAA 362

QY 1547 GAAAAAGTGCTTTGTTTAAATTAATTTGTTGTTGTAATCAATCTGCTTTTCCCAT 1606
Db 361 GAAAAAGTGCTTTGTTTAAATTAATTTGTTGTTGTAATCAATCTGCTTTTCCCAT 302

QY 1607 TGAAGTCTGATTAATCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1666
Db 301 TGAAGTCTGATTAATCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242

QY 1667 CAGCATCTGACAGGTGAATGATGTTCTTCAAGCAATTTCAACCCAGACAGCTGTTTC 1726
Db 241 CAGCATCTGACAGGTGAATGATGTTCTTCAAGCAATTTCAACCCAGACAGCTGTTTC 182

QY 1727 TATCTGTTTAAATAATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1786
Db 181 TATCTGTTTAAATAATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 122

QY 1787 GTCACATAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1846
Db 121 GTCACATAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 62

QY 1847 ATGTGTTTTTGGCAATATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1906
Db 61 ATGTGTTTTTGGCAATATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2

QY 1907 A 1907
Db 1 A 1

RESULT 11
BQ018482/c
DEFINITION
UI-H-DHI-axj-1-02-0-UI-el NCI CGAP_DHI Homo sapiens cDNA clone
IMAGE:5829649 3', mRNA sequence.
ACCESSION
BQ018482
VERSION
BQ018482.1 GI:19753759
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 663)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TUMOR
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
```

```

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
The following repetitive elements were found in this cDNA
sequence: 1-30, >AT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Location/Qualifiers
1..663
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5829649"
/tissue_type="Metastatic Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP_DHI"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP DHI is a normalized cDNA library containing the
following tissue(s): VS-8 Cell line from Metastatic
Chondrosarcoma in Lung. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is AGATCATTCG.
TAG LIB-UI-H-DHI
TAG_TISSUE=Lung
TAG_SEQ=AGATCATTCG"

BASE COUNT      219 a      106 c      140 g      197 t      1 others
ORIGIN
Query Match      30.0%; Score 581.4; DB 12; Length 663;
Best Local Similarity 97.3%; Pred.No. 5.5e-83;
Matches 623; Conservative 0; Mismatches 12; Indels 5; Gaps 3;

QY 1278 GCCAATTTTCAACCCATTTTCTGTGTTTGGGCTCCACATTTGCAATGCAATGCCACGTTG 1337
Db 645 GCCAATTTTCAACCCATTTTCTGTGCT-TGGGCTCCACATTTGCAATGCAATGCCACGTTG 587

QY 1338 CTGCTGACACCGACCGGAGTACTAGCCAGACAAAG-GCAGGTTAGCTGTAATGCTTTT 1396
Db 586 CTGCTGACACCGACCGGAGTACTAGCCAGACAAAGNGCAGGGTAGCCTGAATTTGCTTT 527

QY 1397 CTGCTCTTTTACATTTCTTTTAAATAAGCATTTAGTGTCTCAGTCCCTACTGAGTACTCTT 1456
Db 526 CTGCTCTTTTACATTTCTTTTAAATAAGCATTTAGTGTCTCAGTCCCTACTGAGTACTCTT 467

QY 1457 TCTCTCCCTCTCTCTGAATTTAAATTTCTTCACTTTGCAATTTGCAAGGATTACACATTTTC 1516
Db 466 TCTCTCCCTCTCTCTGAATTTAAATTTCTTCACTTTGCAATTTGCAAGGATTACACATTTTC 407

QY 1517 ACTGTGATGTATATTTGTTTGCAGNAGAAAGAAAAGTGTCTTTGTTTAAATTTACTTTG 1576
Db 406 ACTGTGATGTATATTTGTTTGTGCA---AAAAAAGTGTCTTTGTTTAAATTTACTTTG 350

QY 1577 GTTTGTGAATCCATCTGCTTTTCCCATTTGGACTAGTCTATTAACCCATCTCTGAACT 1636
Db 349 GTTTGTGAATCCATCTGCTTTTCCCATTTGGAACTAGTCTATTAACCCATCTCTGAACT 290

QY 1637 GGTAGAAAAACATCTGAAGAGCTAGTCTATCAGCATCTGACAGGTTGAATTTGATGTTTCT 1696
Db 289 GGTAGAAAAACATCTGAAGAGCTAGTCTATCAGCATCTGACAGGTTGAATTTGATGTTTCT 230

QY 1697 CAGAACCATTTTACCCAGACGCTGTTTCTATCTCTGTTTAAATTTAGTTTGGTTTCT 1756
Db 1697 CAGAACCATTTTACCCAGACGCTGTTTCTATCTCTGTTTAAATTTAGTTTGGTTTCT 1756
```

```

Db      229  CAGAACCAATTTCAACCCAGACGCTGTTTCTATCCTGTTTAAATAAATTAGTTGGTTCT 170
QY      1757  CTACATGATACAAACCTGCTCAATCTGTCTCATATAAAAGTCTGTGACTTCAAGTTTAA 1816
Db      169  CTACATGATACAAACCTGCTCAATCTGTCTCATATAAAAGTCTGTGACTTCAAGTTTAA 110
QY      1817  GTACGACCCCAACCAACTTATTTTCTATGTTTGTGTTTGTGCAATATGAGTGTGTTG 1876
Db      109  GTACGACCCCAACCAACTTATTTTCTATGTTTGTGTTTGTGCAATATGAGTGTGTTG 50
QY      1877  AAAATAAGTACCCATGCTTTTATTAATAAANAATAAATAA 1916
Db      49  AAAATAAGTACCCATGCTTTTATTAATAAANAATAAATAA 10

RESULT 12
LOCUS   BE973555
DEFINITION 601680918F1 NIH_MGC_83 681 bp mRNA linear EST 04-OCT-2000
mRNA sequence.
ACCESSION BE973555
VERSION BE973555
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 681)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lln.gov
Plate: LNC818 row: d column: 05
High quality sequence stop: 614.
Location/Qualifiers
1..681
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3950932"
/lab_host="NIH MGC 83"
/clone_lib="NIH MGC 83"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatggcc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATTCAGAGCGGCGGCGGACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 198 a 144 c 147 g 192 t
ORIGIN

Query Match 28.6%; Score 554.2; DB 10; Length 681;
Best Local Similarity 96.5%; Pred. No. 1.2e-78;
Matches 599; Conservative 0; Mismatches 18; Indels 4; Gaps 3;

QY 849 CCCTGGGGAGCCAGTCCACTGAGCAAGCAAGGAGTGTGAGCTTTTCAGGAAAAG 908
Db 1 CCCTGGGGAGCCAGTCCACTGAGCAAGCAAGGAGTGTGAGCTTTTCAGGAAAAG 60
QY 909 GCTAAGAAAAGAAAACCAATTTCTAAACACAAAGAAACTGTCCAAATGCTTTGGAA 968

```

```

Db      61  GCTAAGAAAAGAAAACCAATTTCTAAACACAAAGAAACTGTCTCAAAATGCTTTGGAA 120
QY      969  CTGTGTTTATTTGCTTATAATGGGTGCCCAAAATGGGTAACTTAGACTTTCAGAGAGATGA 1028
Db      121  CTGTGTTTATTTGCTTATAATGGGTGCCCAAAATGGGTAACTTAGACTTTCAGAGAGATGA 180
QY      1029  GCAGAGAGCAAGGAGAAATCTGGCTGCTCTTCCATTTTCATTTCTGTATCTCAGGTGAG 1088
Db      181  GCAGAGAGCAAGGAGAAATCTGGCTGCTCTTCCATTTTCATTTCTGTATCTCAGGTGAG 240
QY      1089  CTGCTAGAGGGGAGACATTTAGAAAAAATGAAAAACAAACAATTAATTAATGAGGTAG 1148
Db      241  CTGCTAGAGGGGAGACATTTAGAAAAAATGAAAAACAAACAATTAATTAATGAGGTAG 300
QY      1149  CTGAGGCTTGGAGTCTTTGACTCCTACTTAAATCCGTTTGTAGTGAAGAACCTTTCAA 1208
Db      301  CTGAGGCTTGGAGTCTTTGACTCCTACTTAAATCCGTTTGTAGTGAAGAACCTTTCAA 360
QY      1209  TTTTCTTTTATTAGAGGGCCAGCTTACTGTGTGGTGGCAAAATTTGCCAACAATAAGTTAA 1268
Db      361  TTTTCTTTTATTAGAGGGCCAGCTTACTGTGTGGTGGCAAAATTTGCCAACAATAAGTTAA 420
QY      1269  AGAAAGTTGGCAATTTCAACCCATTTCTGTGTTTGGCTCCACATTTGCAATGTTCAA 1328
Db      421  AGAAAGTTGGCAATTTCAACCCATTTCTGTGTTTGGCTCCACATTTGCAATGTTCAA 480
QY      1329  TGCCACGTGCTGTCTACACACCGGAGTACTAGCCAGCAAGCAAGGAGGTAGCTGA 1388
Db      481  TGCCACGTGCTGTG-TCACACCGGAGTACTAGCCAGCAAGCAAGGAGGTAGCTGA 539
QY      1389  ATTGCTTTCTGCTCTTTTACATTTCTTTAAATAAGCAATTTAGTGTCTGCTCCCTACTGA 1448
Db      540  ATGG- -TTTCTGTCTTTTACATTTCTTTAAATAAGCAATTTAGTGTCTGCT-CTTACTGA 596
QY      1449  GTACTCTTTCTCTCCCTCTCT 1469
Db      597  GTACTCTTTCTCTCCCTCTCT 617

RESULT 13
LOCUS   A1989273
DEFINITION 2 prostate cancer cell line LNCap Homo sapiens EST 02-SEP-2001
ACCESSION A1989273
VERSION A1989273.1 GI:15421021
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 563)
AUTHORS Zhang, J.S. and Smith, D.I.
TITLE High throughput screening for androgen regulated genes in LNCap
cells: Identification of hAG-2 as an androgen regulated gene over
expressed in prostate adenocarcinoma
JOURNAL Unpublished
COMMENT Contact: Zhang, Jin-San
Dept. Pathology and Lab Medicine
Mayo Clinic Cancer Center
200 1st St. SW, Rochester MN 55905, USA
Tel: 507-2660311
Fax: 507-2665193
Email: zhang.jinsan@mayo.edu.
Location/Qualifiers
1..563
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="LNCap"
/note="prostate cancer cell line LNCap"
subtracted cDNA libraries from prostate cancer cell line
LNCap treated with androgen."
FEATURES
source

```

BASE COUNT	187 a	104 c	143 g	129 t
Query Match	28.4%;	Score 550;	DB 9;	Length 563;
Best Local Similarity	99.8%;	Pred. No. 6e-78;	Mismatches 0;	Indels 1; Gaps 1;
Matches 561;	Conservative	0;	Mismatches	0; Indels 1; Gaps 1;
QY	586	T T A A A A A T C A T A G C A T T A A A T C C C A A A T C C T A T T A T T A A A G A C C T G A C A G C T T G A G A A G G T	645	
Db	3	T T A A A A A T C A T A G C A T T A A A T C C C A A A T C C T A T T A T T A A A G A C C T G A C A G C T T G A G A A G G T	62	
QY	646	C A C T A C T G C A T T T A T A G G A C C T T C G T G T T C T C T G T T A C C T T A G G T T G A A G T C T G A C A A T C	705	
Db	63	C A C T A C T G C A T T T A T A G G A C C T T C G T G T T C T C T G T T A C C T T A G G T T G A A G T C T G A C A A T C	122	
QY	706	C T T G A G A A T C T T T G C A T G C A G A G G A G T A A G A G G T A T T G G A T T T T C A C A G A G G A A G A C A	765	
Db	123	C T T G A G A A T C T T T G C A T G C A G A G G A G T A A G A G G T A T T G G A T T T T C A C A G A G G A A G A C A	182	
QY	766	C A C G C G A G A A T A A G A G G C C A G C T T A C T A G A G C T G T C C A G T G G A G G G C T C A T G G T G G G A	825	
Db	183	C A C G C G A G A A T A A G A G G C C A G C T T A C T A G A G C T G T C C A G T G G A G G G C T C A T G G T G G G A	241	
QY	826	C A T G G A A A G A G A G C C A G C T A G C C C T G G G A G C C C A G T C C A C T A G A G C A A G G A C T	885	
Db	242	C A T G G A A A A G A A G C C A G C C T A G C C C T G G G A G C C C A G T C C A C T A G A G C A A G G A C T	301	
QY	886	G A G T G A G C C T T T T G C A G A A A A G G C T A G A A A A A G S A A A A C C A T T C T A A A A C A C A C A A G	945	
Db	302	G A G T G A G C C T T T T G C A G A A A A G G C T A G A A A A A G S A A A A C C A T T C T A A A A C A C A C A A G	361	
QY	946	A A A C T G C C A A T G C T T T G G G A A C T G T G T T A T T G C C T A T A A T G G T C C C C A A A A T G G G T	1005	
Db	362	A A A C T G C C A A T G C T T T G G G A A C T G T G T T A T T G C C T A T A A T G G T C C C C A A A A T G G G T	421	
QY	1006	A A C T A G A C T T C A G A G A A T A G A G A G A G A C A A A G A G A A A T C T G G T G T C C T T C A T T	1065	
Db	422	A A C T A G A C T T C A G A G A A T A G A G A G A G A C A A A G A G A A A T C T G G T G T C C T T C A T T	481	
QY	1066	T T C A T T C T G T A C T C A G T G A G C T G T A G A G G G A G A C A T T A G A A A A A A T G A A A C A A C	1125	
Db	482	T T C A T T C T G T A C T C A G T G A G C T G T A G A G G G A G A C A T T A G A A A A A A T G A A A C A A C	541	
QY	1126	A A A C A A T T A C T A A T G A G G T A C	1147	
Db	542	A A A C A A T T A C T A A T G A G G T A C	563	
RESULT 14	BF679599	565 bp mRNA linear EST 21-DEC-2000		
LOCUS	602154041F1 NIH_MGC_83 Homo sapiens cdna clone IMAGE:4294931 5',			
DEFINITION	mRNA sequence.			
ACCESSION	BF679599			
VERSION	EST.			
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
NIH-MGC http://mgc.nci.nih.gov/.				
National Institutes of Health, Mammalian Gene Collection (MGC)				
Unpublished				
Contact: Robert Strausberg, Ph.D.				
Email: cgapbs@mailto.nih.gov				
Tissue Procurement: CLONETECH Laboratories, Inc.				
cDNA Library Preparation: CLONETECH Laboratories, Inc.				
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)				
DNA Sequencing by: Incyte Genomics, Inc.				
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:				
http://image.llnl.gov				

[illegible]